INEGALITES LOG-SOBOLEV POUR LA LOI D'UNE DIFFUSIONET GRANDES DEVIATIONS POUR DES EDP STOCHASTIQUES

Mathieu Gourcy

N^o d'ordre : D.U. 1716

UNIVERSITÉ BLAISE PASCAL

(U.F.R. Sciences et Technologies)

ÉCOLE DOCTORALE DES SCIENCES FONDAMENTALES
N^o : 514

THÈSE

présentée pour obtenir le grade de

DOCTEUR D'UNIVERSITÉ

(Spécialité : MATHÉMATIQUES APPLIQUÉES)

par

Mathieu GOURCY

Diplomé d'Etudes Approfondies

INÉGALITES LOG-SOBOLEV POUR LA LOI D'UNE DIFFUSION ET GRANDES DÉVIATIONS POUR DES EDP STOCHASTIQUES

Soutenue publiquement le 12 décembre 2006 , devant la commission d'examen

Messieurs	Pierre BERNARD	Examinateur
	Arnaud GUILLIN	Examinateur
	Martin HAIRER	Rapporteur
	Etienne PARDOUX	Rapporteur
	Jean PICARD	Président
	Liming WU	Directeur de thèse

Remerciements

Je n'imagine pas commencer mes remerciements autrement qu'en exprimant ma profonde gratitude à l'égard de Liming Wu, mon directeur de Thèse. Sans sa passion, ses idées et sa disponibilité, je n'en serais pas là aujourd'hui. Il m'a fait découvrir, mieux comprendre et surtout aimer la Recherche.

Je remercie vivement les Professeurs Etienne Pardoux et Martin Hairer pour l'honneur qu'ils me font d'être les rapporteurs de ma thèse, et pour m'avoir aidé à repérer et corriger quelques erreurs dans une version préliminaire de ce manuscrit.

Merci à Jean Picard pour avoir accepté de faire partie du jury. J'en profite aussi pour lui témoigner combien mes participations à l'école de Probabilités de Saint-Flour ont été pour moi des expériences scientifiques et humaines exceptionnelles.

Merci également à Pierre Bernard et à Arnaud Guillin, qui me font le plaisir de participer au jury.

Je remercie tous mes collègues du Laboratoire de Mathématiques de l'Université Blaise Pascal pour leur accueil durant ces quelques années, et leurs enseignements les années précédentes. Un grand merci à Bernard Saramito pour ses conseils depuis la Licence et à Hacène Djellout pour ses tuyaux. Merci vraiment à tous les personnels techniques, à Annick pour les articles, à Damien pour les ordinateurs, à Marie-Paule pour les craies, à Valérie pour les missions, etc.... Et à Noëlle pour tout le reste !

Un coucou à mes camarades qui habitent, passent ou sont passés par ce bureau 1213 : Mamadou, Ingrid, Rémi, Mohammed, Liangzen, Benoît, Haija, Nicolas, Malcom, Béranger, Hamid, Lyse, Vivien, Jérémy ... Un gros Merci à Fabrice qui m'a donné du courage quand j'en manquais, et qui a relu et raturé toute ma thèse !

Si je suis venu à l'université pour faire des Maths, c'est que j'ai eu la chance d'avoir au collège, au Lycée puis en Prépa des profs exceptionnels qui m'ont donné envie d'enseigner : Messieurs Murat, Brun, Canonico et Mesdames Cognard, Brancier et Bernard.

Il n'y a plus de place et pourtant il reste tous les copains qui rendent la vie belle. Ils sont trop importants pour que je prenne le risque d'en oublier : merci à *La Colo*, et pourvu que ça dure !

Maman, Papa, je suis bien content de pouvoir écrire ici que je vous aime, et que je suis fier de la façon dont vous avez accompagné vos deux enfants. Je suis aussi très fier de toi Thomas, et de ton courage. Cette thèse je l'ai faite autant pour toi que grâce à toi. Pépé Henri, Mémé Marie, merci pour votre affection, et aussi pour la cuisine et le jardin ! J'ai une pensée pour Mémé Germaine et Pépé Baptiste, et je suis sûr qu'ils sont eux aussi fiers de leurs petits enfants. Ah la famille ...

A Line,
qui ne sait toujours pas où et quand on va vraiment pouvoir habiter ensemble ;-)

Introduction Générale

Dans l'étude de nombreux systèmes issus des Sciences Physiques, l'hypothèse d'ergodicité, validée par l'expérience, est centrale. L'exemple du phénomène de turbulence est typique : une petite différence dans les conditions initiales induit des solutions très éloignées. Il faut alors utiliser une description statistique pour comprendre le système. C'est en fait après un temps suffisamment long pour qu'il ait oublié sa condition initiale, que le système révèle sa structure intrinsèque. On observe alors les états d'équilibre statistique, la ou les mesures invariantes.

Il n'y a que dans de rares cas très simples que l'on a une expression explicite de la mesure invariante, disons μ. Regarder des moyennes d'observations du système le long de la trajectoire constitue un moyen pour obtenir des informations sur μ. En effet, l'ergodicité donne la convergence

$$\frac{1}{T}\int_0^T f(X_t)dt \to \int f(x)\mu(dx) \tag{0.1}$$

lorsque $T \to \infty$, pour X_t une variable aléatoire décrivant l'état du sytème au temps t, et $f(X_t) \in \mathbb{R}$ une observation. De nombreuses questions de recherche sont liées à cette convergence : existence et unicité, dépendance en les conditions initiales, estimations sur la mesure invariante, sur la vitesse ...

Cette thèse a été élaborée à l'Université Blaise Pascal, sous la direction du Professeur Liming WU. Elle contient des résultats nouveaux, classés en deux parties distinctes, et concernant le type de problème évoqué ci-dessus. Les références précises des travaux cités dans cette introduction sont données dans les parties correspondantes.

L'approche développée dans la première partie de cette thèse est d'évaluer le côté droit dans (0.1) par l'espérance du côté gauche à un temps T fixé suffisamment grand. Comme l'information disponible consiste en l'observation d'une trajectoire, il est important de contrôler, pour $r > 0$, la probabilité

$$\mathbb{P}\left(\frac{1}{T}\int_0^T f(X_t)dt > \frac{1}{T}\int_0^T \mathbb{E}f(X_t)dt + r\right) \tag{0.2}$$

de déviation d'une observation à sa moyenne. Les méthodes utilisées pour traiter ce type de problème de concentration sont souvent basées sur certaines inégalités fonctionnelles au niveau trajectoriel (i.e. sous $\mathbb{P}$). En particulier, via l'argument de Herbst, les inégalités de Sobolev logarithmiques (ISL) constituent un outil puissant.

On s'intéresse d'abord au cas général d'un mouvement Brownien avec dérive X_t dans $\mathbb{R}^d$. Les calculs sur le Γ_2 développés par Bakry et Emery permettent d'obtenir une ISL pour la loi de X_t en un temps t fixé, et pour la mesure invariante μ. Mais pour obtenir la concentration de type (0.2), il faut ici une inégalité pour la loi du processus sur l'espace des trajectoires $W^d := C([0,T],\mathbb{R}^d)$. Fang a d'abord établi une inégalité de Poincaré sur W^d relativement à la métrique de Cameron-Martin associée au calcul de Malliavin. Puis des inégalités de Sobolev logarithmiques pour cette métrique ont été prouvées par Hsu, Aida-Elworthy et Capitaine-Hsu-Ledoux.

Cependant, les résultats existants ne répondent pas à notre question de manière satisfaisante. En effet, par rapport à la métrique de Cameron-Martin, les fonctionnelles du type $F(\gamma) = 1/T \int_0^T f(\gamma_t)dt$ ont une constante de Lipschitz qui explose avec le temps. Ceci empêche d'obtenir une borne efficace pour (0.2) lorsque T est grand.

Pour éviter ce phénomène, on a établi une ISL sur l'espace des trajectoires muni de la métrique $L^2([0,T],\mathbb{R}^d)$. Cette inégalité implique bien une borne supérieure pour (0.2) efficace pour les temps T grands, car la constante obtenue pour l'ISL et la constante de Lipschitz de la fonction F sont alors bornées. On étend ensuite le résultat au cas d'une diffusion elliptique générale sur une variété riemannienne, en établissant une nouvelle version de la formule de représentation des martingales de Fang-Clark-Ocone-Haussmann. Dans ces deux cas, l'hypothèse requise est affaiblie : il s'agit d'une minoration sur la courbure de Bakry-Emery, au lieu de sa bornitude dans les résultats connus. Par exemple, pour une diffusion elliptique X avec une courbure minorée par $K > 0$, et si $g \in C_b^1$, on obtient l'inégalité de concentration suivante

$$\mathbb{P}\left(\left| \frac{1}{T}\int_0^T g(X_t)dt - \frac{1}{T}\int_0^T \mathbb{E}g(X_t)dt \right| > r \right) \leq \exp\left(-\frac{r^2K^2T}{2\|\nabla_x g\|_\infty^2} \right).$$

Pour présenter le problème étudié dans la seconde partie de cette thèse, réécrivons (0.1) sous la forme

$$\int f dL_t \to \int f(x)\mu(dx) \tag{0.3}$$

en faisant apparaître la mesure d'occupation $L_T(A) := \int_0^T \delta_{X_t}(A)dt$ qui est une mesure aléatoire sur H si le processus X_t vit dans l'espace H. La question qui se pose est alors celle de la vitesse de convergence de L_T vers μ. Une réponse peut être apportée à l'aide de la théorie des grandes déviations, comme Wu l'a récemment expliqué avec l'étude d'une équation différentielle stochastique modélisant un système Hamiltonien.

Dans ce travail, on s'intéresse aux équations de Navier-Stokes stochastiques régissant le mouvement d'un fluide visqueux, homogène et incompressible en dimension 2

$$dX(t) + \nu AX(t)dt + B\left(X(t), X(t)\right)dt = fdt + GdW(t)\ ;\ X(0) = x. \tag{0.4}$$

Dans l'expression précédente, $\nu > 0$ est la viscosité du fluide, l'opérateur A représente le processus de dissipation, B est le processus de convection et f une force extérieure. Pour schématiser, disons que la perturbation aléatoire prend la forme $GdW(t) = \sum_{k=1}^{\infty} \sigma_k e_k d\beta_k(t)$ où $(e_k)_{k\geq 1}$ est une base de l'espace H des champs de vecteurs L^2 dans lequel vit la solution de (0.4) , les $\beta_k(t)$ sont des mouvements Browniens réels standards, et $\sigma_k \in \mathbb{R}$ est l'amplitude correspondante. Bensoussan-Temam ont d'abord démontré l'existence de solution pour des bruits très réguliers. Flandoli a généralisé ce résultat et établi l'existence d'une mesure invariante μ.

Le problème de l'ergodicité et de la vitesse de convergence vers μ pour cette équation est, depuis quelques années, l'objet d'intenses recherches . L'unicité de la mesure invariante a d'abord été prouvée par Flandoli-Maslowski pour des bruits non dégénérés (avec $\sigma_k \geq 1/\sqrt{k}, \forall k$), et par E-Mattingly-Sinai pour un nombre de modes excités suffisamment grand par rapport à la viscosité ($\sigma_k \neq 0, \forall k \leq N_\nu$). Dans ce cas, la convergence exponentielle $P_t \to \mu$ des mesures de transition a été établie par Bricmont-Kupiainen-Lefevere, Kuksin-Shirikyan et Mattingly grâce à des techniques de couplage. Plus récemment, Hairer-Mattingly-Pardoux ont supprimé cette dépendance en la viscosité et ont prouvé l'ergodicité pour des bruits très dégénérés. Dans ce travail nous continuons l'étude de bruits non dégénérés, pour lesquels Ferrario a prouvé la propriété forte de Feller, l'irréductibilité topologique et donc l'unicité de la mesure invariante, tandis que Goldys-Maslowski ont récemment établi la convergence exponentielle $P_t \to \mu$.

L'objectif de la seconde partie est d'établir un Principe de Grandes Déviations (PGD) pour la mesure d'occupation de la solution de l'équation (0.4). A notre connaissance, l'étude sur les grandes déviations en temps grands pour une équation aux dérivées partielles stochastique était un sujet vierge. A la place d'une convergence macroscopique $P_t \to \mu$, nous obtenons ainsi la convergence exponentielle de la mesure d'occupation $L_t \to \mu$, c'est-à-dire pour la trajectoire. Notre PGD est relatif à la topologie forte de la convergence contre les fonctions mesurables et bornées. Il fournit des estimations asymptotiques en T du type $\mathbb{P}(L_T \in G) \sim e^{-Tc}$ lorsque $T \to \infty$ et $\mu \notin G \subset M_1(H)$, l'espace des mesures de probabilité sur H. Contrairement aux résultats existants, l'exposant c ci-dessus est explicitement connu comme l'infimum de la fonction de taux sur un certain ensemble. De plus, la borne inférieure de grandes déviations donne une minoration de la probabilité de déviation que l'on n'obtient pas avec les autres méthodes.

Si le bruit satisfait à des conditions de non dégénérescence du type $c/k^{1-\varepsilon} \leq \sigma_k \leq C/k^{\frac{1}{2}+\varepsilon}$, alors le semigroupe de transition du processus de Markov $X(t)$ solution de l'équation (0.4) possède la propriété forte de Feller, et d'irréductibilité topologique dans un certain espace fonctionnel $D(A^\alpha)$ associé à l'opérateur A. Comme H, l'espace $D(A^\alpha)$ est de dimension infinie, ce qui rend difficile la caractérisation des compacts. On établit alors des estimations exponentielles a priori sur la solution de (0.4), pour une norme suffisamment forte. En combinant certains résultats généraux de grandes déviations avec ces nouvelles estimations, on peut obtenir le PGD dans l'espace $D\left(A^{\alpha}\right)$. Puis, les propriétés fines des solutions et des arguments topologiques permettent de l'étendre à l'espace H. Enfin, on montre d'autres PGD pour certaines quantités, fonctionnelles de la solution, qui présentent un intérêt physique dans l'étude du phénomène de turbulence, par exemple pour la famille de lois $\mathbb{P}\left(1/T\int_0^T |X(t)|_V dt \ \in \ \cdot\right)_{T>0}$ sur $\mathbb{R}$.

Dans un deuxième temps, on s'intéresse à l'équation (0.4) perturbée par un bruit qui n'est plus blanc en temps, mais une impulsion aléatoire à chaque pas de temps fixe : le second membre de (0.4) est alors remplacé par $\sum_{k\geq 0}\delta(t-k)\eta_k$. Kuksin-Shirikyan, Masmoudi-Young et Shirikyan ont prouvé l'unicité de la mesure invariante et l'ergodicité exponentielle sous différentes conditions sur la structure spectrale de l'impulsion aléatoire $\eta_k \in H$. Ici on s'intéresse juste à l'existence d'une mesure d'équilibre. Pour cela, on établit une nouvelle version d'un critère d'existence sous une condition de dérive à la Meyn-Tweedie. Puis, on montre que cette condition est satisfaite sous l'hypothèse faible que l'amplitude de l'impulsion est log-intégrable, i.e. dès que $\mathbb{E}\log(1+c_0|\eta_k|) < \infty$ pour un $c_0 > 0$.

Enfin, on établit un PGD pour une équation de Burgers stochastique en dimension 1. De nombreux auteurs ont suggéré d'utiliser cette équation comme modèle simple pour la turbulence. Da Prato-Debussche-Temam ont établi l'existence et l'unicité des solutions, et l'existence d'une mesure invariante. Da Prato-Gatarek et Da Prato-Zabczyk ont d'abord prouvé l'unicité de la mesure invariante pour des bruits cylindriques. Nous poursuivons ici l'étude dans le cas de bruits non-dégénérés de trace finie pour lesquels Da Prato-Debussche sont les premiers à avoir obtenu l'unicité, et Goldys-Maslowski la convergence exponentielle. En fait, nous présentons un travail antérieur au cas de Navier-Stokes qui utilise des méthodes analogues dans un cadre un peu plus simple.

Présentation des chapitres

Cette thèse comporte deux parties respectivement introduites par les chapitres 1 et 3.

La première partie, concernant les ISL et leur utilisation pour obtenir la concentration de certaines fonctionnelles de la trajectoire, est composée des deux premiers chapitres. Dans le premier chapitre on introduit en fait différentes notions concernant les ISL et certains objets de géométrie riemannienne. De plus, on motive et on présente les résultats principaux prouvés dans le chapitre 2 ainsi qu'une variante concernant des systèmes dynamiques en temps discret. Le deuxième chapitre est un article en collaboration avec Liming Wu publié dans la revue Potential Analysis qui constitue le coeur de cette partie.

La seconde partie comporte les chapitres 3,4,5 et 6 concernant le comportement asymptotique des équations stochastiques de Navier-Stokes bidimensionnelle et Burgers unidimensionnelle. Dans le troisième chapitre, on introduit plus précisément les équations de Navier-Stokes, et on motive l'ajout d'une force aléatoire. Puis on rappelle quelques résultats connus sur le problème stochastique associé, avant de présenter le principe de grandes déviations ainsi que les résultats démontrés dans les chapitres 4,5 et 6 rédigés en anglais. Le chapitre 4 est un article accepté par la revue Stochastic Processes and their Applications, qui contient les preuves des résultats nouveaux concernant le comportement ergodique des équations de Navier-Stokes. Le chapitre 5 est consacré à l'existence d'une mesure invariante dans le cas d'un bruit formé d'impulsions (ou "kicks"). Enfin, le chapitre 6 contient un PGD pour l'équation de Burgers stochastique. Cet article rédigé par l'auteur est accepté pour publication aux Annales de l'Institut Henri Poincaré.

On peut considérer le chapitre 7 comme une annexe à ce document. Il contient des résultats connus de grandes déviations, essentiellement dus à Wu. Pour faciliter l'utilisation de ce document, ils sont ici présentés sous les hypothèses et notations des précédents chapitres. Certains éléments de preuves sont rappelés.

L'auteur espère que le choix de faire se succéder dans ce document des chapitres de synthèse en français et des chapitres en anglais sous forme conventionnelle de (pré)publications permettra une lecture fluide, et s'excuse de certaines répétitions qui en résultent.

Sommaire

Première partie

Inégalités de Sobolev logarithmiques pour la loi d'une diffusion

Chapitre 1

Inégalités de Sobolev logarithmiques sur l'espace des trajectoires

Dans cette partie, nous présentons l'étude de l'inégalité de Sobolev logarithmique sur l'espace des trajectoires muni de la métrique L^2. Ce choix de métrique est motivé par l'obtention de propriétés de concentration pour certaines fonctionnelles de la trajectoire qui restent pertinentes pour les temps T grands, contrairement à ce que l'on pouvait avoir avec les résultats existants.

La fin de ce chapitre est consacrée à un problème indépendant. On s'intéresse à des systèmes dynamiques en temps discret type $X_{n+1} = F(X_n, W_{n+1}) \in E$ où $(W_n)_{n\geq 1}$ est un suite de v.a.i.i.d modélisant une perturbation aléatoire. On cherche alors des conditions pour que la loi du processus $(X_n)_{n\geq 0}$ sur $E^{\mathbb{N}}$ satisfasse à une inégalité de Poincaré ou de Sobolev logarithmique.

1.1 Introduction et motivation du changement de métrique

Rappelons tout d'abord quelques définitions dans un cadre assez général. Soit $(X, \mathcal{B}, \mu)$ un espace de probabilité. On convient de noter $\mathbb{E}^\mu$ l'intégration contre la mesure de probabilité μ et $(L^p(\mu), \|\cdot\|_p)$ les espaces de Lebesgue sur $(X, \mathcal{B}, \mu)$.

Pour une fonction f de carré intégrable par rapport à la mesure μ, on notera par la suite

$$Var_\mu(f) = \mathbb{E}^\mu(f^2) - (\mathbb{E}^\mu(f))^2$$

la variance de f. Si maintenant f est une fonction positive et intégrable par rapport a la mesure μ, on définit de la manière suivante l'entropie de f par rapport a μ

$$Ent_\mu(f) = \mathbb{E}^\mu(f \log f) - \mathbb{E}^\mu(f) \log \mathbb{E}^\mu(f).$$

Comme l'application $x \to x \log x$ est convexe et minorée sur $\mathbb{R}^+$, et avec l'inégalité de Jensen, la condition f intégrable garantie l'existence de l'entropie dans $[0, +\infty]$. En fait l'entropie de f par rapport à μ est finie si et seulement si $\mathbb{E}^\mu \left(f \log^+ f\right) < \infty$.

On dit qu'une mesure de probabilité μ satisfait à une Inégalité de Sobolev Logarithmique (ISL) si il existe une constante $\alpha(\mu) > 0$ telle que pour toute fonction f assez régulière l'inégalité suivante est satisfaite :

$$Ent_\mu(f^2) \leq 2\alpha(\mu)\mathbb{E}^\mu|\nabla f|^2.$$

C'est Gross [17] qui introduit cette inégalité en 1975, l'établit d'abord pour la mesure de Bernoulli, et l'obtient par tensorisation et avec le théorème de la limite centrale pour la mesure gaussienne avec une constante optimale $\alpha(\mu)$ égale à 2. L'inégalité de Sobolev logarithmique précédente implique celle de Poincaré

$$Var_\mu(f) \leq \alpha(\mu)\mathbb{E}^\mu|\nabla f|^2.$$

Les champs d'utilisation des ISL sont nombreux. On peut citer les travaux [6] de Bakry-Emery sur leur fameux critère Γ_2, Bakry [5] et Ledoux [21] pour les études systématiques et applications, Bobkov-Götze [7], Otto-Villani [23], Bobkov-Gentil-Ledoux [8, 16], Djellout-Guillin-Wu [13] pour les liens avec les inégalités de transport. Dans le travail qui suit, on s'intéresse particulièrement aux liens avec certains problèmes de concentration de loi du processus. Dans [21], Ledoux présente une synthèse de ces questions, reprenant entre autres des travaux de Aida, Masuda, Shigekawa et Stroock [1, 3, 4]. On établit ici, à l'aide d'une nouvelle ISL, de bonnes propriétés de concentration pour certaines fonctionnelles de la trajectoire d'une diffusion.

Laissons donc le cadre général pour présenter une ISL sur un espace des trajectoires. Dans le cas du mouvement Brownien dans $\mathbb{R}^d$, on peut aussi attribuer à Gross [17] le théorème suivant. Soit μ la mesure de Wiener sur $W^d = C\big([0, T], \mathbb{R}^d\big)$ et $h \in H$, l'espace de Cameron Martin, i.e, l'espace de Hilbert des fonctions absolument continues $h : [0, T] \to \mathbb{R}^d$ vérifiant $h(0) = 0$ et

$$\|h\|_H^2 := \int_0^T |\dot{h}(t)|^2 dt < \infty.$$

Théorème 1.1.1. *Pour toute fonction f sur W^d régulière et bornée*

$$Ent_\mu(f^2) \leq 2C(T)\mathbb{E}^\mu \left(\|Df\|_H^2\right) \tag{1.1}$$

pour la constante optimale $C(T) = 1$.

Ici, $Df(\gamma) = (Df(\gamma,t))_{t\in[0,T]}$ est le gradient de Malliavin dans H, dont la construction est par exemple détaillée dans le livre de Nualart [22]. Rappelons que la dérivée directionnelle selon h d'une fonction régulière f en un $\gamma \in W^d$ est

$$D_h f(\gamma) = \lim_{\varepsilon\to 0} \frac{f(\gamma+\varepsilon h) - f(\gamma)}{\varepsilon} \tag{1.2}$$

et que l'on a alors

$$\forall h \in H, D_h f(\gamma) = \langle Df(\gamma), h\rangle_H = \int_0^T \dot{\widehat{D}}f(\gamma,t)\cdot \dot{h}(t)dt \tag{1.3}$$

où $\dot{\widehat{D}}f(\gamma,t) := \frac{d}{dt}Df(\gamma,t)$.

On présente d'abord, en suivant [21], une méthode appelée argument de Herbst qui permet de déduire des propriétés de concentration pour une mesure satisfaisant à une inégalité de Sobolev logarithmique. Le but est de souligner les propriétés requises sur la fonctionnelle F et sur l'inégalité de Sobolev logarithmique vérifiée par une mesure ν sur W^d pour obtenir une propriété de concentration efficace en temps T grand.

Théorème 1.1.2. *On suppose que ν vérifie l'inégalité de Sobolev logarithmique*

$$Ent_\nu(f^2) \le 2C(T)\mathbb{E}^\nu|\nabla f|^2$$

et soit $F : W^d \to \mathbb{R}$ une fonction régulière telle que pour toute trajectoire $\gamma \in W^d$,

$$|\nabla F(\gamma)|^2 \le \sigma^2(T).$$

Alors on a la concentration quadratique en r suivante

$$\nu\left(F \ \ge\ \mathbb{E}^\nu F + r\right) \le e^{-\frac{r^2}{2C(T)\sigma^2(T)}}.$$

Démonstration. Pour un énoncé et une preuve complète, il faudrait préciser entre autres des propriétés de la classe de fonctions sur laquelle l'inégalité est vérifiée, ainsi que des propriétés sur le gradient ∇ utilisé (celles que l'on attend d'une dérivation, par exemple une règle de composition permettant le calcul (1.4) ci dessous). De telles propriétés sont vérifiées pour le gradient de Malliavin, et le gradient L^2 introduit par la suite. Pour une preuve complète dans le cadre général, on renvoie au livre de Ledoux [21].

On rappelle ici l'argument central sur la transformée de Laplace de F sous la mesure ν, définie par

$$H(\lambda) = \mathbb{E}^\nu\left(e^{\lambda F}\right).$$

Si, pour tout $\lambda > 0$, on peut appliquer l'inégalité de Sobolev logarithmique à la fonction $f^2 = e^{\lambda F}$, on obtient alors

$$\begin{aligned}\lambda H'(\lambda) - H(\lambda)\log H(\lambda) &\leq 2C(T)\frac{\lambda^2}{4}\mathbb{E}^\nu\left(e^{\lambda F}|\nabla F|^2\right)\\ &\leq 2C(T)\frac{\lambda^2}{4}\sigma^2(T)H(\lambda)\end{aligned}$$

puisque l'on sait que

$$|\nabla f|^2 = |\nabla e^{\frac{\lambda}{2}F}|^2 = \frac{\lambda^2}{4}e^{\lambda F}|\nabla F|^2 \tag{1.4}$$

et que $|\nabla F| \leq \sigma^2(T)$. De plus, comme la transformée de Laplace est strictement positive, on peut réécrire l'inégalité précédente sous la forme

$$\frac{H'(\lambda)}{\lambda H(\lambda)} - \frac{\log H(\lambda)}{\lambda^2} \leq \frac{2C(T)\sigma^2(T)}{4},$$

la quantité à gauche de l'inégalité précédente n'étant autre que la dérivée $K'(\lambda)$ de la fonction

$$K(\lambda) = \frac{1}{\lambda}\log H(\lambda).$$

Remarquant la convergence de $K(\lambda)$ vers $\mathbb{E}^\nu F$ lorsque λ tend vers 0, ainsi que le fait que K est continue sur $[0,\lambda]$ et dérivable dans l'intérieur de cet intervalle, la formule de Newton-Leibniz s'écrit

$$K(\lambda) - K(0) = \int_0^\lambda K'(s)ds \leq \frac{2C(T)\sigma^2(T)}{4}\lambda.$$

Si on traduit cette estimation pour la transformée de Laplace, on obtient

$$H(\lambda) \leq e^{\lambda\mathbb{E}^\nu F + \frac{2C(T)\sigma^2(T)}{4}\lambda^2}.$$

Il suffit alors de remarquer que l'inégalité de Markov donne

$$\nu\left(F \ \geq \ \mathbb{E}^\nu F + r\right) = \nu\left(e^{\lambda F} \ \geq \ e^{\lambda(\mathbb{E}^\nu F + r)}\right) \leq e^{-\lambda r + \lambda^2\frac{2C(T)\sigma^2(T)}{4}}$$

et de conclure grâce à une optimisation en $\lambda > 0$. □

Pour que l'argument de Herbst donne des résultats de concentration intéressants en temps grands pour une fonctionnelle F donnée, il apparaît donc deux conditions nécessaires :
(i) la constante $C(T)$ dans l'inégalité de Sobolev logarithmique ne doit pas exploser en temps T.
(ii) De même, la borne supérieure $\sigma(T)$ pour la constante de Lipschitz de la fonctionnelle F ne doit pas exploser en temps T.

Par exemple, pour la mesure de Wiener dans $\mathbb{R}^d$, la constante $C(T)$ du Théorème 1.1.1 n'explose pas. Cependant, l'extension sur une variété riemannienne de courbure de Ricci uniformément bornée par K prouvée par Capitaine, Hsu et Ledoux [10] fait intervenir la constante $C(T) = e^{KT}$ et ne vérifie donc pas la condition (i). On va voir qu'avec la métrique de Cameron-Martin, même si la constante $C(T)$ n'explose pas (comme par exemple dans le travail de Aida-Elworthy [2]), c'est la constante $\sigma(T)$ qui explose.

On s'intéresse ici à certaines fonctionnelles F de la diffusion qui devraient posséder de bonnes propriétés de concentration si la diffusion a un *bon* comportement ergodique. C'est le cas par exemple de la fonctionnelle associée au théorème de la limite centrale

$$F_T(\gamma) = \frac{1}{\sqrt{T}} \int_0^T (g(\gamma_t) - \mathbb{E}^\mu g(\gamma_t))\, dt \tag{1.5}$$

avec $g : \mathbb{R}^d \to \mathbb{R}$ régulière.

Or, si la métrique considérée sur l'espace des trajectoires est celle de Cameron-Martin, comme dans le Théorème 1.1.1, la condition (ii) n'est malheureusement pas vérifiée par la fonction (1.5) même sous l'hypothèse suivante de majoration du supremum de la norme du gradient (usuel) de g en x,

$$\|\nabla_x g\|_\infty := \sup_{x\in\mathbb{R}^d} |\nabla_x g(x)| < \infty.$$

En effet, un calcul direct sur la dérivée directionnelle (1.2) donne

$$\dot{\widehat{DF_T}}(\gamma, t) = \frac{1}{\sqrt{T}} \int_t^T \nabla_x g(\gamma_s) ds$$

et on obtient la majoration

$$\|DF_T\|_H^2 \leq \frac{\|\nabla_x g\|_\infty^2 T^2}{3}$$

dont le côté droit tend vers $+\infty$ avec T. On observe encore ce phénomène, même avec une renormalisation type loi des grands nombres, i.e pour les fonctionnelles du type

$$G_T(\gamma) = \frac{1}{T} \int_0^T (g(\gamma_t) - \mathbb{E}^\mu g(\gamma_t))\, dt.$$

Pour obtenir des résultats intéressants en temps grands, la métrique L^2 sur l'espace des trajectoires semble mieux adaptée. Un exemple simple permet de s'en convaincre : regardons le cas où $d = 1$ et $g(x) = x$ dans (1.5), et considérons deux trajectoires continues $\gamma_0(t) = 0$ et $\gamma_1(t) = t$. On a $\|\gamma_1 - \gamma_0\|_H = \sqrt{T}$, et

$$|F_T(\gamma_1) - F_T(\gamma_0)| = \frac{T^2}{2\sqrt{T}} = \frac{T}{2}\|\gamma_1 - \gamma_0\|_H$$

ce qui ne permet pas d'espérer une bonne constante de Lipschitz, alors que pour la norme $L^2 := L^2([0,T],\mathbb{R}^d)$, on a $\|\gamma_1 - \gamma_0\|_{L^2} = \frac{T^{\frac{3}{2}}}{\sqrt{3}}$ qui est du bon ordre.

Djellout, Guillin et Wu [13] ont obtenu une inégalité de transport sur cet espace L^2 pour les diffusions elliptiques à valeurs dans $\mathbb{R}^d$. La section suivante est une suite à ce travail en établissant l'inégalité de Sobolev logarithmique pour la loi d'un mouvement Brownien avec dérive sur l'espace des trajectoires muni de la métrique L^2, d'abord dans le cas plat puis sur une variété riemannienne.

1.2 Mouvement Brownien avec dérive dans $\mathbb{R}^d$

Soit W_t un mouvement Brownien standard dans $\mathbb{R}^d$, x un point initial dans $\mathbb{R}^d$ et une dérive $b : \mathbb{R}^d \to \mathbb{R}^d$ continuement différentiable. On considère $X_t(x)$ la diffusion dans $\mathbb{R}^d$ définie par l'équation différentielle stochastique

$$dX_t = dW_t + b(X_t)dt \ , \quad X_0 = x \in \mathbb{R}^d. \tag{1.6}$$

On suppose de plus que la dérive b vérifie

$$(\nabla b)^\sigma \le -K \ Id, \ \ K \in \mathbb{R} \tag{1.7}$$

où Id est la matrice identité dans $\mathbb{R}^d$, $\nabla b = (\partial_j b_i)_{ij}$ la matrice jacobienne, A^σ est la matrice symétrisée $(A + A^t)/2$, et $A \le B$ s'entend au sens des matrices symétriques positives.

On reconnaît la condition (1.7) de courbure dimension de Bakry-Emery [5, 6] qui implique une ISL pour la loi de X_t en un $t > 0$ fixé, et si $K > 0$ pour l'unique mesure invariante m du processus de Markov (X_t).

Nous nous intéressons ici à la loi du processus sur l'espace $L^2([0,T],\mathbb{R}^d)$ des trajectoires. Pour introduire notre gradient, nous considérons donc $W^d = C([0,T],\mathbb{R}^d)$ en tant que sous espace dense de $L^2 := L^2([0,T],\mathbb{R}^d)$, et nous adoptons la notion usuelle de différentiabilité sur un espace vectoriel normé.

Définition 1.2.1. *Soit $f : W^d \subset L^2 \to \mathbb{R}$ une fonction différentiable pour la norme L^2 au sens classique. On note $\nabla f(\gamma) := (\nabla_t f(\gamma))_{t\in[0,T]}$ l'élément de $L^2([0,T],\mathbb{R}^d)$ vérifiant*

$$D_g f(\gamma) = \langle \nabla f(\gamma) \ , \ g\rangle_{L^2([0,T],\mathbb{R}^d)} \ , \ \forall g \in H.$$

Sa relation avec le gradient de Malliavin (1.3) est donnée par le lemme suivant. Ce résultat est en particulier utile pour donner une définition indirecte d'un gradient L^2 dans le cas des diffusions sur une variété riemannienne M.

Lemme 1.2.2. *Si $f : W^d \to \mathbb{R}$ est différentiable pour la norme L^2, alors son gradient de Malliavin Df existe. De plus, pour toute trajectoire $\gamma \in W^d$, $\overset{\frown}{\ddot{D}}f(\gamma)$ est dans $L^2([0,T],\mathbb{R}^d)$ au sens des distributions et*

$$\nabla_t f(\gamma) = -\overset{\frown}{\ddot{D}}f(\gamma,t), \quad dt - p.p.$$

On définit enfin $C_b^1(W^d/L^2)$ l'espace des fonctions f bornées sur W^d, qui sont différentiables pour la norme $L^2([0,T],\mathbb{R}^d)$, et telles que leur gradient ∇f soit aussi une fonction continue et bornée de $(W^d, \|\cdot\|_{L^2})$ dans $L^2([0,T],\mathbb{R}^d)$. On peut alors énoncer notre premier résultat : une ISL pour la loi de X.

Théorème 1.2.3. *Soit $\mathbb{P}_x$ la loi de la diffusion $X.(x)$ définie par l'EDS (1.6). Si b vérifie (1.7), alors pour toute fonction $f \in C_b^1(W^d/L^2)$,*

$$Ent_{\mathbb{P}_x}(f^2) \leq 2C(T)\mathbb{E}^{\mathbb{P}_x}\int_0^T |\nabla_t f|^2 dt \tag{1.8}$$

pour

$$C(T) := \begin{cases} \frac{1}{K^2}\left(1-\sqrt{2e^{-KT}-e^{-2KT}}\right), & si\ \ 2e^{-KT}-e^{-2KT}>0, K\neq 0 \\ \frac{T^2}{2}, & si\ \ K=0 \\ \frac{1}{K^2}\left(e^{-2KT}-2e^{-KT}+1\right), & si\ \ 2e^{-KT}-e^{-2KT}\leq 0, K\neq 0. \end{cases} \tag{1.9}$$

En particulier, pour $K>0$, on a $C(T) \leq 1/K^2$ (pour $T>0$).

Par exemple, dans le cas du processus de Ornstein-Uhlenbeck unidimensionnel

$$dX_t = dW_t - \frac{X_t}{2}dt,\ X_0 = 0$$

on constate que la constante $C(T) \to 4$ est asymptotiquement optimale, pour $T \to \infty$. On sait en effet qu'une mesure Gaussienne sur l'espace de Hilbert séparable $G := L^2([0,T],\mathbb{R})$ vérifie une ISL avec une constante optimale $\lambda_{max}(\Sigma)$ plus grande valeur propre de Σ, matrice de covariance. En dimension infinie, Σ est un opérateur de Hilbert-Schmidt. Or pour le processus Gaussien X_t à trajectoires dans G, son opérateur de covariance est donné par $\Sigma : f \in L^2 \to \Sigma f \in L^2$ où

$$\Sigma f(s) = \int_0^T cov(X_s, X_t) f(t) dt$$

et en particulier pour le processus de Ornstein-Uhlenbeck, on peut calculer

$$cov(X_s,X_t) = \exp\left(-\frac{|t-s|}{2}\right) - \exp\left(\frac{s+t}{2}\right).$$

On en déduit que
$$\lambda_{max}(\Sigma) \geq \frac{1}{T}\langle \Sigma 1_{[0,T]}, 1_{[0,T]}\rangle_{L^2}$$
où la quantité de droite tend vers 4 quand $T \to \infty$.

Remarquons aussi que le lemme 1.2.2 permet une comparaison entre l'énergie disons *de Malliavin* et l'énergie L^2 sur un intervalle borné $[0,T]$. On a en effet

$$\mathbb{E}^{\mathbb{P}_x}\|Df\|_H^2 \leq \frac{T^2}{2}\mathbb{E}^{\mathbb{P}_x}\int_0^T |\nabla_t f|^2 dt,$$

ce qui permet de retrouver directement le cas $K = 0$. Cependant, pour le cas ergodique intéressant $K > 0$, le même raisonnement aboutit à une constante explosive contrairement à celle fournie par notre Théorème 1.2.3.

La preuve de ce théorème est directe : nous avons montré que l'application

$$\Phi : (W^d, \|\cdot\|_H) \to (W^d, \|\cdot\|_{L^2})$$

qui, à une trajectoire Brownienne W, fait correspondre une trajectoire X de la diffusion est Lipschitzienne. Cette approche *solution forte*, ne fonctionne plus dès que l'on ajoute un coefficient de volatilité $\sigma(X_t)$ non constant devant le terme dW_t dans (1.6). Cependant, ce dernier cas peut parfois être ramené à celui des diffusions elliptiques à valeurs dans une variété riemannienne M, c'est pourquoi nous étendons le Théorème 1.2.3 à ce nouveau cadre par une méthode complètement différente.

1.3 Une brève introduction géométrique

Le but de cette section est d'introduire quelques notions nécessaires de géométrie riemannienne et de géométrie stochastique. Les bases de la géométrie différentielle sont supposées connues (cartes locales, espaces et fibrés tangents, ...), et on ne présente que quelques points classiques. On suit le livre de Boothby [9] pour le cadre riemannien et les travaux de Hsu [18] pour la partie stochastique.

1.3.1 De la géométrie riemannienne...

Sur une variété M de dimension d, une métrique riemannienne est la donnée en tout point x de M d'un produit scalaire $ds_x^2 = (\cdot|\cdot)_x$ sur T_xM, l'espace tangent à la variété M au point x, qui est ainsi isométrique à $\mathbb{R}^d$. Soient $x = (x^i)$ une carte locale et $X_i = \partial/\partial x^i$ une base de T_xM, alors en coordonnées locales la métrique riemannienne s'écrit
$$ds^2 = g_{ij}dx^i dx^j, \quad g_{ij} = (X_i|X_j),$$
où la matrice $g = (g_{ij})$ est définie positive en chaque point. Une variété munie d'une métrique riemannienne est appelée variété riemannienne.

Soit M une variété différentielle. On note TM le fibré tangent et $\Gamma(TM)$ le $C^\infty(M)$-module des champs de vecteurs de classe C^∞ sur M. Il n'existe pas en général d'identification canonique entre T_xM et T_yM, si x et y sont deux points distincts de M, mais par contre il existe un moyen d'identifier T_{γ_0} et T_{γ_t} pour une courbe régulière $\gamma : t \in [0,T] \to M$ à l'aide d'une connexion. Une connexion sur M est une manière de différentier un champ de vecteurs le long d'un autre champ de vecteurs. C'est donc une application

$$\nabla : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM),$$

qui vérifie de plus les propriétés suivantes : pour tous $X, Y, Z \in \Gamma(TM)$ et $f, g \in C^\infty(M)$,

$$\begin{array}{l} 1)\nabla_{fX+gY}Z = f\nabla_X Z + g\nabla_Y Z \\ 2)\nabla_X(Y+Z) = \nabla_X Y + \nabla_X Z \\ 3)\nabla_X(fY) = f\nabla_X Y + X(f)Y. \end{array}$$

$\nabla_X Y$ est appelée dérivée covariante de Y le long de X.

Dans le cas d'une variété riemannienne M, il existe une unique connexion Γ, appelée connexion de Levi-Civita, qui est compatible avec la métrique riemannienne, i.e.

$$\forall X, Y, Z \in \Gamma(TM), \quad Z\left((X|Y)\right) = (\nabla_Z X, Y) + (X, \nabla_Z Y)$$

et qui de plus est sans torsion (ou symétrique), i.e.

$$\forall X, Y \in \Gamma(TM), \quad \nabla_X Y - \nabla_Y X = [X, Y] \tag{1.10}$$

où $[X, Y]$ est le crochet de Lie des champs de vecteurs X et Y. En coordonnées locales, une connexion peut s'exprimer à l'aide des symboles de Christoffel, qui sont connus explicitement dans le cas de la connexion de Levi-Civita. Dans la suite la variété riemannienne M est munie de cette connexion.

Un propriété fondamentale des variétés riemanniennes est que l'on peut définir la notion de courbure, qui généralise la notion usuelle de courbure pour des surfaces de l'espace. Pour tous champs de vecteurs X, Y, Z, W, le tenseur riemannien de courbure R_m en ces vecteurs est défini par

$$R_m(X, Y, Z, W) = \left(\nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z | W\right).$$

Ce tenseur est donc assez simple à définir, mais il est peu intuitif et lourd à manipuler dans les calculs. C'est pourquoi, à partir de ce tenseur, on définit la courbure sectionnelle d'un plan σ contenu dans l'espace tangent T_xM, et ayant pour base orthonormée (X, Y) :

$$K(\sigma) = K(\{\{X, Y\}\}) = R_m(X, Y, Y, X).$$

Pour tout point x dans M, et tout plan σ de T_xM, $K_x(\sigma)$ est un nombre : c'est la courbure de Gauss au point x de la sous-variété $\exp_x(\sigma)$ de M.

On peut alors introduire la notion de courbure de Ricci comme moyenne de courbures sectionnelles : Soit X un vecteur normal dans T_xM, et $(X, X_2, ..., X_d)$ une base orthonormale de T_xM, alors $Ric_x(X,X)$ est définie par

$$Ric_x(X,X) = \sum_{k=2}^{d} K(\{\{X, X_k\}\}).$$

En fait, cette moyenne des $K(\sigma)$ pour tous les plans σ de T_xM générés par $\{X, X_j\}$ ne dépend pas du choix de $(X_2, \dots, X_d)$ et pour tout point x, la courbure de Ricci Ric_x est une forme quadratique sur l'espace tangeant T_xM. Pour la suite, on peut donc retenir qu'il s'agit d'une matrice symétrique Ric_u, si un repère orthonormé u de T_xM est fixé.

Un champ de vecteurs (V_t) le long d'une courbe (x_t) de M est dit parallèle le long de cette courbe si on a $\nabla_{\dot{x}} V = 0$ en tout point de la courbe. Dans ce cas, le vecteur V_{x_t} au point x_t est le transport parallèle de V_{x_0} le long de cette courbe. Un champ de vecteurs parallèle V le long d'une courbe (x_t) est entièrement déterminé par sa valeur initiale V_{x_0}.

Une courbe (x_t) de M est appelée une géodésique si $\nabla_{\dot{x}}\dot{x} = 0$ le long de (x_t), c'est-à-dire si le champ des vecteurs vitesses est lui-même parallèle le long de la courbe. Une géodésique est donc complètement déterminée par sa position et sa vitesse initiales.

Le produit scalaire permet aussi de définir une distance sur M ; nous supposerons dans la suite que M muni de cette distance est un espace métrique complet. (D'après le théorème de Hopf et Rinow, M est alors géodésiquement complète, i.e. tout segment géodésique peut être étendu à l'infini des deux côtés, et tout couple de points de M peut être connecté par une géodésique minimisant la distance.)

La fin de cette partie est consacrée à la notion de champ de vecteurs horizontal, afin de pouvoir donner par la suite une construction du mouvement Brownien sur M. Considérons $O(M)$, le fibré des repères orthonormés (dont chaque fibre $O(M)_x$ est constituée des isométries de $\mathbb{R}^d$ dans T_xM), et la projection canonique $\pi : O(M) \to M$. La connexion ∇ permet de scinder l'espace tangent au repère $u \in O(M)$ en un espace vertical et un espace horizontal :

$$T_uO(M) = V_uO(M) \oplus H_uO(M).$$

Le sous-espace $V_uO(M)$ est l'espace des vecteurs tangents verticaux, c'est-à-dire tangents à la fibre $O(M)_{\pi(u)}$. Pour la définition de $H_uO(M)$, considérons une courbe (u_t) dans $O(M)$, c'est-à-dire un repère se promenant de manière C^∞ le long de la courbe $\pi(u_t)$ de M. Cette courbe (u_t) est dite horizontale si pour tout vecteur e de $\mathbb{R}^d$, le champ de vecteurs $(u_t(e))$ est parallèle le long de $\pi(u_t)$; un vecteur

$\xi \in T_uO(M)$ est dit horizontal si c'est le vecteur tangent d'une courbe horizontale partant de u. L'ensemble de ces vecteurs horizontaux en u forme l'espace vectoriel $H_uO(M)$.

La projection canonique π induit alors en tout $u \in O(M)$ un isomorphisme linéaire

$$\pi_{*u} : H_uO(M) \to T_{\pi(u)}M$$

qui permet de relever un vecteur tangent $v \in T_{\pi(u)}M$ en un vecteur tangent horizontal, noté v^*. Si $(e_1, \dots, e_d)$ désigne la base canonique de $\mathbb{R}^d$, on définit pour tout $u \in O(M)$

$$H_i(u) = H_{e_i}(u) = (u(e_i))^*.$$

On obtient ainsi les d champs de vecteurs horizontaux fondamentaux sur $O(M)$. Ils engendrent en chaque point u l'espace $H_uO(M)$.

Soit (u_t) un relèvement horizontal d'une courbe (x_t) de M (c'est-à-dire que (u_t) est une courbe horizontale et $\pi(u_t) = x_t$, u_0 étant fixé). Comme $\dot{x}_t \in T_{x_t}M$, on a $u_t^{-1}\dot{x}_t \in \mathbb{R}^d$. L'antidéveloppement de la courbe (x_t) (ou de la courbe horizontale (u_t)) est une courbe dans $\mathbb{R}^d$ définie par

$$w_t = \int_0^t u_s^{-1}\dot{x}_s ds. \tag{1.11}$$

Remarquons que w dépend du choix du repère initial u_0 de manière très simple : si (v_t) est un autre relèvement horizontal de (x_t) et $u_0 = v_0 g$ pour un certain $g \in O_d(\mathbb{R})$, alors l'antidéveloppement de (v_t) est (gw_t). De plus, comme $u_t\dot{w}_t = \dot{x}_t$, on a

$$H_{\dot{w}_t}(u_t) = (u_t(\dot{w}_t))^* = (\dot{x}_t)^* = \dot{u}_t.$$

Et finalement l'antidéveloppement (w_t) et le relèvement horizontal (u_t) d'une courbe (x_t) de M sont reliés par l'équation différentielle ordinaire dans $O(M)$:

$$\dot{u}_t = \sum_{i=1}^d H_i(u_t)\dot{w}_t^i. \tag{1.12}$$

Réciproquement, si on part d'une courbe (w_t) de $\mathbb{R}^d$ et d'un repère orthonormé u_0 en un point x_0 d'origine, l'équation (1.12) admet une unique solution (u_t) qui est une courbe horizontale de $O(M)$. C'est le développement de la courbe (w_t) dans $O(M)$, et sa projection $(x_t) = (\pi u_t)$ est appelée développement de (w_t) dans M. La méthode qui permet de passer de (w_t) à (x_t) est connue sous le nom de *roulement sans glissement*.

Cette construction permet de définir de manière intrinsèque le mouvement Brownien, et donc la mesure de Wiener μ_M, sur une variété riemannienne M de dimension d.

1.3.2 ... à l'analyse stochastique.

Le Laplacien usuel sur M est l'opérateur de Laplace-Beltrami Δ_M défini avec la métrique riemannienne $g = g_{ij}dx^i dx^j$ en coordonnées locales par la formule

$$\Delta_M f = \frac{1}{\sqrt{\det g}} \frac{d}{dx^i} \left(\sqrt{\det g} g^{ij} \frac{df}{dx^j} \right)$$

où g^{ij} est la matrice inverse de g_{ij}. On peut construire un mouvement Brownien sur M directement à partir de Δ_M, mais il n'y a pas de façon intrinsèque (à la variété M) pour écrire Δ_M sous la forme d'une somme de carrés de champs de vecteur. Il faut pour cela considérer la variété M comme sous variété de l'espace euclidien $\mathbb{R}^l$ (théorème de Nash) avec l souvent strictement plus grand que d. Dans ce cas, le mouvement Brownien euclidien conduisant l'EDS définissant celui sur M contient trop d'informations.

Nous suivons dans ce paragraphe la construction de Eells-Elworthy-Malliavin exposée par Hsu [19].

La construction commence avec $O(M)$, le fibré des repères orthonormés de M. On rappelle que $\pi : O(M) \to M$ est la projection canonique, et que $H_1, ..., H_d$ sont les d champs de vecteurs horizontaux canoniques sur M. L'opérateur sous-elliptique du second ordre sur $O(M)$

$$\Delta^H = \sum_{i=1}^{d} H_i^2$$

est appelé Laplacien horizontal de Bochner.

En fait cet opérateur est relié à Δ_M par la formule suivante valide pour toute fonction f de classe C^2 sur M et pour tout repère orthonormé $u \in O(M)$

$$\Delta^H (f \circ \pi)(u) = \Delta_M f(\pi u).$$

Comme Δ^H est une somme de d carrés et un relèvement de Δ_M, on peut construire un mouvement Brownien sur $O(M)$ en résolvant une EDS définie sur $O(M)$ conduite par $W = (W^i)$ un mouvement Brownien sur $\mathbb{R}^d$. On fixe un point x_0 de M et un repère orthonormé u_0 en ce point. On considère alors l'équation différentielle stochastique

$$dU_t = \sum_{i=1}^{d} H_i(U_t) \circ dW_t^i \ , \ U_0 = u_0 \in O(M). \tag{1.13}$$

Sa solution U est une diffusion sur $O(M)$ de générateur $\frac{\Delta^H}{2}$ que l'on appelle mouvement Brownien horizontal. En appliquant la formule de Itô pour U, et en projetant sur M, on peut montrer grâce à la formule reliant Δ^H à Δ_M que $X = \pi U$

est une diffusion sur M de générateur $\frac{\Delta_M}{2}$. Ainsi X est un mouvement Brownien sur M issu de $x_0 \in M$. Dans la suite, on notera μ_M sa loi sur

$$W_{x_0}(M) = \{f : [0,T] \to M \text{ continues t.q. } f(0) = x_0\}.$$

On peut en effet prouver que le mouvement Brownien sur M, diffusion de générateur $\frac{\Delta_M}{2}$, est unique en loi en faisant la construction inverse, c'est à dire en montrant qu'il peut être relevé en un mouvement Brownien horizontal, i.e. en une diffusion de générateur $\frac{\Delta^H}{2}$. Il s'agit pour cela de résoudre l'équation différentielle stochastique en U sur $O(M)$ suivante

$$dU_s = \sum_{i=1}^{d} H_i(U_s)(U_s^{-1} \circ dX_s)^i. \tag{1.14}$$

La résolution est assez technique [19], il faut d'abord la réécrire comme une EDS de Stratonovich sur $O(M)$. Cette solution est appelée relèvement horizontal de X. En fait, comme dans le cas déterministe, on peut montrer que le processus stochastique (B_t) à valeurs dans $\mathbb{R}^d$ défini par

$$B_t = \int_0^t U_s^{-1} \circ dX_s$$

est un mouvement Brownien sur l'espace euclidien appelé développement stochastique de (X_t), ce qui explique le lien entre les équations (1.13) et 1.14). Le relèvement horizontal et le développement stochastique de X dépendent tous les deux du choix d'un repère orthonormé initial.

1.4 Diffusion elliptique sur une variété

On commence par énoncer le résultat pour la loi μ_M du mouvement Brownien sur M.

1.4.1 Le cas du mouvement Brownien dans M

Des inégalités de Sobolev logarithmiques pour la loi μ_M du mouvement Brownien sur une variété riemannienne complète de courbure de Ricci bornée ont été établies par Aida-Elworthy [2], Hsu [19] et Capitaine-Hsu-Ledoux [10]. Dans ces travaux, le gradient sur l'espace des trajectoires est celui de Malliavin, la métrique celle de Cameron-Martin. Aida-Elworthy ont obtenu une constante $C(T)$ qui n'explose pas en utilisant l'injection isométrique $M \subset \mathbb{R}^l$, et le résultat dans le cas plat (Théorème 1.1.1). D'après le Théorème 1.2.3 pour la métrique L^2, la constante obtenue pour le mouvement Brownien dans $\mathbb{R}^d$ est $C(T) = T^2/2$. Le résultat que nous obtiendrions avec cette méthode ne serait pas intéressant pour T grand.

Nous suivons Capitaine-Hsu-Ledoux [10]. Rappelons que, comme dans le cas euclidien, un opérateur gradient de Malliavin existe. Il est d'abord défini pour les fonctions cylindriques $f : \gamma \in W_{x_0}(M) \to f(\gamma_{t_1}, \gamma_{t_2}, ..., \gamma_{t_n}) \in \mathbb{R}$, par

$$Df(X,t) := \sum_{i=1}^{n} U_{t_i}^{-1} \frac{df}{dx_i}(\cdot) t_i \wedge t \in \mathbb{R}^d.$$

Puis à l'aide d'une formule d'intégration par partie, on montre que cet opérateur est fermable sous la condition géométrique de courbure de Ricci bornée

$$\sup_{u \in O(M)} \|Ric_u\| < \infty, \tag{1.15}$$

la différence avec le cas euclidien est en effet l'apparition d'un terme de courbure dans la formule d'IPP de Bismut-Driver.

On note alors $(D, \mathbb{D}_2(D))$ sa fermeture dans $L^2(\mu_M)$. Ainsi pour $h \in H$, $D_h f(\gamma)$ et $Df(\gamma) \in H$ vérifient :

$$\forall h \in H,\ D_h f(\gamma) = \langle Df(\gamma)\ ,\ h \rangle_H = \int_0^T \widehat{\dot{D}f}(\gamma, s) \cdot \dot{h}(s) ds.$$

Dans ce cadre, Capitaine-Hsu-Ledoux [10] ont obtenu l'inégalité de Sobolev logarithmique suivante :

Théorème 1.4.1. *Pour toute fonction F dans le domaine $\mathbb{D}_2(D)$ du gradient*

$$\mathbb{E}(F^2 \log F^2) - \mathbb{E}(F^2) \log \mathbb{E}(F^2) \leq 2C_{CHL}(T) \mathbb{E}(\|DF\|_H^2)$$

avec $C_{CHL}(T) = e^{\sup_{u \in O(M)} \|Ric_u\|\ T}$.

Leur preuve est basée sur une formule de représentation des martingales appliquée à $F(X)$ pour les fonctions $F \in \mathbb{D}_2(D)$. Afin d'utiliser une méthode analogue, on introduit directement notre gradient L^2 sur un sous ensemble de $\mathbb{D}_2(D)$, via le Lemme 1.2.2.

Définition 1.4.2. *Pour $F \in \mathbb{D}_2(D)$ vérifiant $\widehat{\ddot{D}F} \in L^2([0,T], \mathbb{R}^d)$ on définit le gradient de F pour la norme L^2 par*

$$\nabla_t F(\gamma) = -\widehat{\ddot{D}F}(\gamma, t) \quad dt - p.p.$$

Ainsi $\nabla F(\gamma) \in L^2([0,T],\mathbb{R}^d)$ est déterminé par

$$\langle \nabla F(\gamma)\ ,\ h\rangle_{L^2} \ = \ D_h F(\gamma)\ ,\ \forall h \in C_0^\infty(]0,T[,\mathbb{R}^d).$$

Comme on l'a souligné dans le cas plat, le contrôle de l'énergie *de Malliavin* par l'énergie L^2 implique, avec le Théorème , une ISL de constante $\frac{T^2}{2}C_{CHL}(T)$ inexploitable pour T grand.

Pour obtenir une constante correcte on établit une version de la formule de représentation de Fang, Clark, Ocone, Haussmann [10, 15] pour notre gradient ∇.

Lemme 1.4.3. *On suppose que M vérifie (1.15). Pour $F \in \mathbb{D}_2(D)$ vérifiant $\mu_M -$ p.p, $\overset{\ddot{\frown}}{D}F(\gamma) \in L^2([0,T],\mathbb{R}^d)$ et $\overset{\dot{\frown}}{D}F(\gamma,T) = 0$, on a*

$$F(X) = \mathbb{E}\big(F(X)\big) + \int_0^T \langle H_t\ , dW_t\rangle \tag{1.16}$$

avec

$$H_t = \mathbb{E}\left(\int_t^T A_t^* A_s^{*-1}\nabla_s F ds|\mathcal{B}_t\right).$$

Dans cette expression, $(\mathcal{B}_t)$ est la filtration naturelle de (W_t), et $(A_t)_{0\le t\le T}$ le processus matriciel défini par

$$\frac{dA_t}{dt} - \frac{1}{2}A_t Ric_{U_t} = 0 \quad , \quad A_0 = Id. \tag{1.17}$$

On peut alors démontrer le

Théorème 1.4.4. *On suppose la courbure de Ricci uniformément bornée (1.15), ainsi que la condition de type Bakry-Emery*

$$\forall u \in O(M)\ ,\ \ \frac{1}{2}Ric_u \ge KId,\ \ K \in \mathbb{R}. \tag{1.18}$$

Alors pour $F \in \mathbb{D}_2(D)$ vérifiant $\mu_M -$ p.p, $\overset{\ddot{\frown}}{D}F(\gamma) \in L^2([0,T],\mathbb{R}^d)$ et $\overset{\dot{\frown}}{D}F(\gamma,T) = 0$, on a

$$Var_{\mu_M}\big(F\big) \le C(T)\mathbb{E}^{\mu_M}\int_0^T |\nabla_t F|^2 dt$$

et

$$Ent_{\mu_M}\big(F^2\big) \le 2C(T)\mathbb{E}^{\mu_M}\int_0^T |\nabla_t F|^2 dt$$

avec la constante $C(T)$ définie par (1.9).

La géométrie est prise en compte : on remarque que si $K > 0$, et donc M compacte, alors la constante $C(T)$ dans l'ISL pour la loi μ_M du MB est bornée contrairement au cas du MB dans $\mathbb{R}^d$.

L'expression de $C(T)$ en K est donnée par (1.9) comme dans le cas plat, alors que K n'a pas la même signification. Ceci provient de l'utilisation dans les deux preuves d'un lemme technique identique de contrôle d'une intégrale réelle, mais dans des cadres relativement différents.

On remarque aussi que la constante $C(T)$ ne dépend que de la borne inférieure sur la courbure de Ricci : cette condition (1.18) suffit à établir un lemme type Gronwall matriciel différent de celui utilisé par Capitaine-Hsu-Ledoux. En utilisant ce nouvel outil dans leur preuve, on obtient d'ailleurs une constante non explosive à la place de $C_{CHL}(T)$.

On peut en fait se passer de l'hypothèse de courbure de Ricci uniformément bornée (1.15), qui intervient pour définir $\mathbb{D}_2(D)$, en donnant une définition directe de notre gradient sur une classe restreinte de fonctions test suffisamment riche pour nos applications à la concentration. De plus, le résultat de la section suivante étend le Théorème 1.4.4 à la loi de la diffusion de générateur $\frac{\Delta_M}{2} + V$ dans M.

1.4.2 La diffusion de générateur $\frac{\Delta_M}{2} + V$ dans M

Capitaine, Hsu et Ledoux [10] ont établi une ISL relativement à la métrique de Cameron-Martin pour la loi de la diffusion de générateur $L = \frac{1}{2}\Delta + V$ dans M équipée de n'importe quelle connexion compatible avec la métrique riemannienne dont la torsion vérifie certaines conditions. On continue ici avec la connexion usuelle vérifiant (1.10).

Soit V un champ de vecteur régulier sur M. On va prouver une ISL pour la loi μ_V sur $W_{x_0}(M)$ de la diffusion de générateur $\frac{\Delta}{2} + V$. On suppose que le M.B n'est pas explosif (ainsi $\mu_V \ll \mu_M$) et que

$$\frac{1}{2}Ric_u - (\nabla V)^\sigma \geq KId, \ K \in \mathbb{R}. \tag{1.19}$$

Cette condition de minoration de la courbure de Bakry-Emery couvre les deux situations précédentes : le cas euclidien si $Ric_u = 0$ et le cas de M.B dans M si $V = 0$.

Pour pouvoir se passer de l'hypothèse

$$\sup_{u \in O(M)} \|Ric_u\| < \infty$$

on doit donc restreindre la classe sur laquelle on définit notre ∇, la définition 1.4.2 ne convenant plus.

Définition 1.4.5. *Soit*

$$\mathcal{D} = \left\{ \; F(\gamma) \;\; = \;\; \Phi\left(\int_0^T f_1(s,\gamma_s)ds \;\; , \;\; \ldots \;\; , \;\; \int_0^T f_m(s,\gamma_s)ds\right) \; \right\}$$

avec $m \in \mathbb{N}^*$, $\Phi \in \mathcal{C}_b^\infty(\mathbb{R}^m, \mathbb{R})$ *et* $f_i \in \mathcal{C}_0^\infty([0,T] \times M)$.

On définit sur $\mathcal{D}$ *un gradient* ∇ *pour la norme* L^2 *par*

$$\nabla_t \int_0^T f(s,\gamma_s)ds = U_t^{-1} \nabla f(t,\gamma_t)$$

et la règle de composition

$$\nabla_t \Phi\left(\int_0^T f_1(s,\gamma_s)ds \;\; , \;\; \ldots \;\; , \;\; \int_0^T f_m(s,\gamma_s)ds\right) := \sum_i \partial_i \Phi(\cdots) U_t^{-1} \nabla f_i(t,\gamma_t).$$

Les fonctionnelles du type (1.5) dont on s'intéresse aux propriétés de concentration sont bien dans la classe $\mathcal{D}$. Ce gradient correspond à celui de la partie précédente si $\|Ric_u\|$ est bornée : on peut alors calculer pour la fonction (1.5) précédente

$$\dot{\widehat{DF}}_T(\gamma,t) = \frac{1}{\sqrt{T}} \int_t^T U_s^{-1} \nabla_x g(\gamma_s) ds$$

ce qui nous donne $\nabla_t F_T(\gamma) = \frac{1}{\sqrt{T}} U_t^{-1} \nabla_x g(\gamma_t)$ pour la définition 1.4.2, soit le même résultat qu'en suivant les règles ci dessus définissant ∇ sur la classe $\mathcal{D}$.

Grâce à la propriété de support compact des fonctions de $\mathcal{D}$, et par une technique de localisation on peut établir un résultat analogue au lemme 1.4.3, pour les fonctions de $\mathcal{D}$, et sous la seule condition (1.19). Ce qui nous permet de montrer le

Théorème 1.4.6. *Si* (X, M) *verifie la condition de Bakry-Emery (1.19), alors pour* $F \in \mathcal{D}$ *on a*

$$Var_{\mu^V}(F) \leq C(T) \mathbb{E}^{\mu^V} \int_0^T |\nabla_t F|^2 dt,$$

et

$$Ent_{\mu^V}(F^2) \leq 2C(T) \mathbb{E}^{\mu^V} \int_0^T |\nabla_t F|^2 dt,$$

où la constante $C(T)$ *est définie par (1.9) .*

1.5 Applications à la concentration

Donnons maintenant deux applications immédiates des résultats obtenus. Soit une variété M, et un champ de vecteurs V vérifiant les hypothèses du théorème 1.4.6 pour un $K > 0$. On commence par l'inégalité de concentration de type Hoeffding, la principale motivation de ce travail.

On trouve une inégalité de concentration sans le terme parasite en T qu'on aurait obtenu avec la métrique de Cameron-Martin, bien que le résultat ci dessous ne permette d'ailleurs pas de *lire* la métrique choisie sur l'espace des trajectoires.

Proposition 1.5.1. *Soit $T > 0$, $g \in C_0^\infty([0,T] \times M)$ et soit F_T la fonctionnelle :*

$$F_T(\gamma) = \frac{1}{\sqrt{T}} \int_0^T (g(t,\gamma_t) - \mathbb{E}g(t,X_t))\, dt.$$

Alors

$$\mu_V\left(F_T(X) > r\right) \le \exp\left(-\frac{r^2K^2}{2\|\nabla_x g\|_\infty^2}\right),\ \forall r > 0. \tag{1.20}$$

En effet, comme $\exp(\frac{\lambda F_T}{2}) \in \mathcal{D}$, que $C(T) \le 1/K^2$ pour $K > 0$ et que

$$|\nabla F_T|_{L^2([0,T],\mathbb{R}^d)} \le \|\nabla_x g\|_\infty,$$

l'ISL implique l'estimation suivante qui permet de prouver la proposition

$$\mathbb{E}^{\mu_V} e^{\lambda F_T(X)} \le \exp\left(\frac{\lambda^2\|\nabla_x g\|_\infty^2}{2K^2}\right),\ \forall \lambda \in \mathbb{R}.$$

Bobkov, Gentil et Ledoux ont montré comment obtenir une inégalité de Tsirel'son à partir d'une ISL. Cette méthode combinée à notre ISL donne la propriété suivante sur la fonctionnelle $F_{T,\delta}$ liée au principe d'invariance.

Proposition 1.5.2. *Pour la fonctionnelle*

$$F_{T,\delta}(\gamma) := \frac{1}{\sqrt{T}} \sup_{0\le s\le t\le s+\delta\le 1} \int_{Ts}^{Tt} (g(u,\gamma_u) - \mathbb{E}g(u,X_u))du \tag{1.21}$$

avec $g \in C_b^1([0,T] \times M)$, on a l'inégalité de concentration

$$\mathbb{E}e^{\lambda F_{T,\delta}(X)} \le \exp\left(\lambda \mathbb{E}F_{T,\delta}(X) + \frac{\delta\lambda^2\|\nabla V\|_\infty^2}{2K^2}\right),\ \forall \lambda \in \mathbb{R};$$

$$\mathbb{P}\left(F_{T,\delta}(X) - \mathbb{E}F_{T,\delta}(X) > r\right) \le \exp\left(-\frac{r^2K^2}{2\delta\|\nabla_x g\|_\infty^2}\right),\ \forall r > 0.$$

La partie suivante est consacrée à quelques remarques sur l'analogue en temps discret de l'espace des trajectoires, et sur l'inégalité de Sobolev logarithmique dans ce cadre.

1.6 Remarques sur quelques systèmes dynamiques indexés par $\mathbb{N}$

Soit (E, d_E) un espace métrique muni de sa tribu borélienne $\mathcal{B}$. On suppose que pour tout $x \in E$ et tout $r > 0$, la boule $B(x,r)$ ne contient pas que l'élément x. On peut alors introduire sur E une généralisation naturelle du module du gradient définie par

$$|\nabla f(x)| := \limsup_{d_E(x,y)\to 0} \frac{|f(x)-f(y)|}{d_E(x,y)} \tag{1.22}$$

pour une fonction $f : E \to \mathbb{R}$. Ainsi pour une mesure de probabilité μ sur E, la quantité

$$\mathbb{E}^\mu |\nabla f|^2$$

a un sens au moins pour f lipschitzienne et bornée. L'opérateur ∇ défini par (1.22) possède les propriétés d'une dérivation. En particulier, pour une fonction $\psi : \mathbb{R} \to \mathbb{R}$ de classe $\mathcal{C}^\infty$, on a

$$|\nabla \psi(f)| = |\nabla f||\psi'(f)|$$

et donc

$$\mathbb{E}^\mu |\nabla \psi(f)|^2 \leq \|\nabla f\|_\infty^2 \mathbb{E}^\mu |\psi'(f)^2|.$$

On peut parler d'ISL dans ce cadre [21].

On s'intéresse ici à la loi du système dynamique non linéaire perturbé aléatoirement

$$X_0 = x \in E, \quad X_{n+1}(x) = F\left(X_n(x), W_{n+1}\right), \quad n \geq 0. \tag{1.23}$$

Le bruit $(W_n)_{n\geq 0}$ est une suite de variables aléatoires indépendantes et identiquement distribuées (vaiid), définies sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, et prenant leurs valeurs dans (G, d_G) un autre espace métrique. On notera $\mathbb{E}$, Var et Ent la moyenne, la variance et l'entropie sous la probabilité $\mathbb{P}$.

Enfin, la fonction $F(x,w) : E \times G \to E$ est aussi lipschitzienne bornée par rapport à la variable w. Plus précisément, on suppose que, $\forall (x,w) \in E \times G$,

$$\limsup_{d_G(w,w_1)\to 0} \frac{d_E\left(F(x,w), F(x,w_1)\right)}{d_G(w,w_1)} := |\nabla_w F(x,w)| \leq C \tag{1.24}$$

pour une constante $C > 0$ indépendante de (x,w).

Sous cette condition, et si la loi de W_1 notée $\mathcal{L}(W_1)$ vérifie une inégalité de Poincaré ou de Sobolev logarithmique, alors la probabilité de transition donnant la loi de X_{n+1} sachant celle de X_n aussi.

Lemme 1.6.1. *i) Si $\mathcal{L}(W_1)$ satisfait à l'inégalité de Poincaré sur G,*

$$Var\left(f(W_1)\right) \leq A\, \mathbb{E}\left(|\nabla f(W_1)|^2\right) \tag{1.25}$$

pour une constante $A > 0$, alors pour tout $x \in E$, la loi $\mathcal{L}(X_1(x)) = P(x, \cdot)$ de $F(x, W_1)$ satisfait à une inégalité analogue sur E de constante AC^2.

ii) De même, si la loi du bruit $\mathcal{L}(W_1)$ vérifie l'ISL

$$Ent\left(f^2(W_1)\right) \leq 2\ A\ \mathbb{E}\left(|\nabla f(W_1)|^2\right) \tag{1.26}$$

alors pour tout $x \in E$, la loi $P(x, \cdot)$ satisfait a une ISL sur E de constante AC^2.

Démonstration. Avec les propriétés du gradient et la régularité de F, la preuve est immédiate. On a en effet

$$\begin{aligned} Ent\left(f^2\left(X_1(x)\right)\right) &= Ent\left(f \circ F\left(x, W_1\right)^2\right) \\ &\leq 2A\mathbb{E}\left(|\nabla_w f \circ F(x, W_1)|^2\right) \\ &= 2A\mathbb{E}\left(|(\nabla f) \circ F(x, W_1)|^2 \times |\nabla_w F(x, W_1)|^2\right) \\ &\leq 2AC^2\mathbb{E}\left(|\nabla f\left(X_1(x)\right)|^2\right). \end{aligned}$$

□

Le but de cette partie est d'obtenir de telles inégalités au niveau de la loi de la chaîne de Markov (X_n).

Pour une fonction régulière $f(x_1, \ldots, x_n)$ sur E^n, on définit le module du gradient sur l'espace produit par

$$|\nabla f|^2 = \sum_{l=1}^{n} |\nabla_l f|^2$$

où $\nabla_l(f)$ est la dérivée de f considérée comme fonction de x_l, les $n-1$ autres x_i étant fixés. On introduit aussi les itérées de la fonction F données par

$$X_k(x) = \phi_k\left(x, \omega\right)$$

où $\omega \in G^{\mathbb{N}^*}$ est une réalisation de la suite des bruits. On a ainsi $\phi_1 = F$ et

$$X_{k+l}(x) = \phi_l\left(X_k(x), \Theta^k\omega\right)$$

où $\Theta^k\omega = (W_{k+1}, W_{k+2}, \ldots)$ est un opérateur de décalage.

On établit le

Théorème 1.6.2. *Soit (X_n) la chaîne de Markov définie par (1.23), où F vérifie (1.24) pour un $C > 0$, et (W_n) est une suite de v.a.i.i.d de loi $\mathcal{L}(W_1)$.*

(i) Si la loi $\mathcal{L}(W_1)$ vérifie l'inégalité de Poincaré (1.25) de constante $A > 0$ et si

$$\sup_{x \in E} \mathbb{E}|\nabla_x \phi_l(x, \omega)|^2 \leq {b_l}^2 \tag{1.27}$$

avec $B := \sum_{l=1}^{\infty} b_l < \infty$, alors la loi de la chaîne $(X_n)_{n \in \mathbb{N}}$ vérifie

$$Var\Psi\left(X_1(x), \ldots, X_n(x)\right) \leq A\ C^2\ B^2\ \mathbb{E}|\nabla\Psi\left(X_1(x), \ldots, X_n(x)\right)|^2$$

$\forall x \in E,\quad \forall n \in \mathbb{N},\quad \forall \Psi : E^n \to \mathbb{R}$ régulière.

(ii) De plus, si la loi $\mathcal{L}(W_1)$ satisfait à l'ISL (1.26) de constante $A > 0$ et cette fois ci

$$\sup_{x \in E} |\nabla_x \phi_l(x, \omega)|^2 \leq {b_l}^2, \quad \mathbb{P} - p.s. \tag{1.28}$$

avec $B := \sum_{l=1}^{\infty} b_l < \infty$, alors la loi de la chaîne (X_n) vérifie

$$Ent\left(\Psi^2\right) \leq 2\ A\ C^2\ B^2\ \mathbb{E}|\nabla\Psi|^2$$

$\forall x \in E, \quad \forall n \in \mathbb{N}, \quad \forall \Psi : E^n \to \mathbb{R}$ *régulière.*

Les constantes C et A sont données par (1.24) et le Lemme 1.6.1, en particulier C, A et B ne dépendent ni de $x \in E$ ni de $n \in \mathbb{N}$. Regardons quelques exemples simples d'application de ces résultats. On va prendre par la suite $E = G = \mathbb{R}^d$, de sorte que les dérivations introduites précédemment soient les dérivations usuelles.

L'exemple où tout fonctionne est celui d'un système dissipatif avec un bruit additif (i.e. non corrélé à la position du système) défini par

$$X_{n+1} = S(X_n) + W_{n+1}, \tag{1.29}$$

pour une application $S : \mathbb{R}^d \to \mathbb{R}^d$ contractante, c'est à dire telle que

$$\sup_{x \in \mathbb{R}^d} \|\nabla S(x)\| := r < 1,$$

et pour un bruit quelconque dont la loi vérifie l'inégalité de Sobolev logarithmique sur $\mathbb{R}^d$, disons par exemple que W_1 suit une loi gaussienne. On peut prouver à l'aide d'une récurrence immédiate que

$$\sup_{x \in E} |\nabla_x \phi_l(x, \omega)| \leq \left(\sup_{x \in E} \|\nabla S(x)\|\right)^l.$$

Ainsi, les conditions (1.27) et (1.28) sont vérifiées par le système (1.29) avec $b_l = r^l$, qui est sommable.

On perçoit la grande différence entre les deux conditions de type Ljapunov (1.27) et (1.28) dans le cas d'un bruit multiplicatif, i.e. avec un coefficient de diffusion non constant. Considérons en effet le système dynamique dans $\mathbb{R}$ défini par

$$X_{n+1} = s(X_n) + \sigma(X_n) W_{n+1}. \tag{1.30}$$

Pour la condition (1.28), il faut imposer une perturbation W avec un support borné, hypothèse peu acceptable pour les applications. Par contre, il suffit d'avoir $\mathbb{E}|W|^2 < \infty$ pour que certaines fonctions s et σ vérifient (1.27). En effet, on a le

Lemme 1.6.3. *La suite* $u_k := \sup_{x\in E} \mathbb{E}|\nabla_x \phi_k(x,\omega)|^2$ *est sous additive et donc la suite* $(u_k)^{\frac{1}{k}}$ *admet une limite notée* l.
i) Si $l < 1$, *alors* $\sum_{k\geq 0}(u_k)^{\frac{1}{2}} < \infty$. *En particulier, la condition (1.27) est vérifiée pour* $b_k := (u_k)^{\frac{1}{2}}$.
ii) Si il existe un entier k *tel que* $u_k < 1$, *alors la limite vérifie* $l < 1$.

Démonstration. Il suffit de conditionner pour obtenir

$$\begin{aligned}
\mathbb{E}|\nabla_x \phi_{l+k}(x,\omega)|^2 &= \mathbb{E}|\nabla_x \phi_l\left(X_k(x),\Theta^k\omega\right)|^2 \\
&= \mathbb{E}|\nabla_x \phi_l\left(\cdot,\Theta^k\omega\right)\nabla_x X_k(x)|^2 \\
&= \mathbb{E}\left\{|\nabla_x X_k(x)|^2 \mathbb{E}\left(\nabla_x \phi_l\left(\cdot,\Theta^k\omega\right)|^2 \mid \mathcal{F}_k\right)\right\} \\
&= u_l \mathbb{E}|\nabla_x X_k(x)|^2.
\end{aligned}$$

En prenant le supremum pour x dans E, on a $u_{k+l} \leq u_k u_l$. Les propriétés classiques des suites sous additives qui en découlent sont bien connues. □

Ainsi, même avec une perturbation Gaussienne, la condition (1.27) est vérifiée dès que $u_1 < 1$, par exemple si $\|\nabla b\|_\infty \leq \alpha$ et $\|\nabla\sigma\|_\infty \leq \beta/\mathbb{E}|W|^2$, avec $\alpha+\beta < 1/2$.

1.6.1 Preuve des tensorisations (Théorème 1.6.2)

On commence par démontrer le (i). Pour obtenir l'inégalité de Poincaré au niveau du processus, on utilise une méthode martingale assez classique. Soit $n \in \mathbb{N}$ fixé et Ψ une fonction régulière sur E^n. On considère alors la martingale M_n définie par

$$M_k = \mathbb{E}\left[\Psi\left(X_1, X_2, \ldots, X_n\right)|\mathcal{F}_k\right],$$

où

$$\mathcal{F}_k = \sigma\left(X_1, \ldots, X_k\right) \tag{1.31}$$

est la tribu engendrée par les k premiers pas de la chaîne. On a $M_0 = \mathbb{E}\Psi$, $M_n = \Psi$ et

$$M_k = \Psi\left(x_1, \ldots, x_k, \Phi_1\left(x_k,\Theta^k\omega\right), \ldots, \Phi_{n-k}\left(x_k,\Theta^k\omega\right)\right).$$

Les propriétés de la dérivation généralisée impliquent que

$$|\nabla_{x_k} M_k| = \sum_{l=k}^{n} \mathbb{E}\left[|\nabla_l \Psi||\nabla_{x_k}\Phi_{l-k}\left(x_k,\Theta^k\omega\right)|\,|\mathcal{F}_k\right]$$

où $\Phi_0(x,\omega) = x$. Une martingale ayant des accroissements orthogonaux, la variance prend la forme

$$Var(\Psi) = \sum_{k=1}^{n} \mathbb{E}(M_k - M_{k-1})^2$$

De plus, l'inégalité de Poincaré du Lemme 1.6.1 pour la fonction de transition appliquée à la variance conditionnelle $\mathbb{E}(M_k - M_{k-1})^2$ donne

$$\mathbb{E}(M_k - M_{k-1})^2 \leq AC^2 \mathbb{E}|\nabla_{x_k} M_k|^2.$$

Avec l'expression précédente de $|\nabla_{x_k} M_k|$ et l'inégalité de Cauchy-Schwartz, on obtient

$$\begin{aligned} Var(\Psi) &\leq \sum_{k=1}^n AC^2 \mathbb{E} \left| \sum_{l=k}^n \mathbb{E}\left[\nabla_l \Psi \cdot \nabla_{x_k} \Phi_{l-k}\left(x_k, \Theta^k \omega\right) | \mathcal{F}_k\right] \right|^2 \\ &\leq \sum_{k=1}^n AC^2 \mathbb{E} \left(\sum_{l=k}^n \mathbb{E}\left[b_{l-k} |\nabla_l \Psi|^2 | \mathcal{F}_k\right] \sum_{l=k}^n \mathbb{E}\left[\frac{|\nabla_{x_k} \Phi_{l-k}\left(x_k, \Theta^k \omega\right)|^2}{b_{l-k}} | \mathcal{F}_k \right] \right) \end{aligned}$$

où les b_k vérifient (1.27). Ainsi,

$$Var(\Psi) \leq AC^2 \sum_{k=1}^n \mathbb{E} \left(\sum_{l=k}^n \mathbb{E}\left[b_{l-k} |\nabla_l \Psi|^2 | \mathcal{F}_k\right] \times B \right)$$

et on conclut par le théorème de Fubini.

Pour le (ii), on rappelle d'abord qu'une formulation équivalente (voir [21]) à l'ISL pour la probabilité de transition du Lemme 1.6.1 est donnée par

$$Ent \; g\left(X_1(x)\right) \leq \frac{AC^2}{4} \mathbb{E} \left(\frac{\nabla g\left(X_1(x)\right)}{g\left(X_1(x)\right)} \right) \tag{1.32}$$

pour $g : E \to \mathbb{R}_+^*$ régulière.

Soit g une fonction réelle positive sur E^n, et $\mathcal{F}_k$ la tribu définie par (1.31), on considère la martingale

$$M_k = \mathbb{E}\left[g\left(X_1, X_2, \ldots, X_n\right) | \mathcal{F}_k\right],$$

de sorte que $M_0 = \mathbb{E} g\left(X_1, X_2, \ldots, X_n\right)$ et $M_n = g\left(X_1, X_2, \ldots, X_n\right)$. Quitte à renormaliser, on peut supposer $M_0 = \mathbb{E} g = 1$. On a donc

$$\begin{aligned} Ent(g) &= \mathbb{E}\left[M_n \log(M_n)\right] \\ &= \sum_{k=1}^n \mathbb{E}\left[M_k \log\left(\frac{M_k}{M_{k-1}}\right)\right]. \end{aligned}$$

Comme $\mathcal{L}(W_1)$ satisfait à une ISL, le noyau de transition aussi d'après le Lemme 1.6.1. On peut donc majorer l'entropie conditionnelle précédente avec la formulation (1.32) de l'ISL, ce qui nous donne

$$\mathbb{E}\left[M_k \log\left(\frac{M_k}{M_{k-1}}\right)\right] \leq \frac{AC^2}{4} \mathbb{E}\left[\frac{|\nabla_{x_k} M_k|^2}{M_k}\right].$$

Maintenant, pour $g = \Psi^2$ on peut calculer

$$\begin{aligned}
|\nabla_{x_k} M_k| &= \left| \nabla_{x_k} \mathbb{E}[\Psi^2 \left(X_1(x), X_2(x), \ldots, X_n(x)\right) |\mathcal{F}_k] \right| \\
&\leq \sum_{l=k}^{n} \mathbb{E}\left[2|\Psi|\ |\nabla_l \Psi|\ |\nabla_{x_k} \Phi_{l-k}\left(x_k, \Theta^k \omega\right)|\ \ |\mathcal{F}_k\right].
\end{aligned}$$

Ainsi avec la condition (1.28), et une majoration analogue à celle utilisée par Capitaine, Hsu et Ledoux [10], on obtient

$$\begin{aligned}
|\nabla_{x_k} M_k|^2 &\leq 4\mathbb{E}\left(\sum_{l=k}^{n} |\Psi|^2 \frac{|\nabla_{x_k} \Phi_{l-k}\left(x_k, \Theta^k \omega\right)|^2}{b_{l-k}} \mid \mathcal{F}_k\right) \times \mathbb{E}\left(\sum_{l=k}^{n} |\nabla_l \Psi|^2 b_{l-k} \mid \mathcal{F}_k\right) \\
&\leq 4\mathbb{E}\left(|\Psi|^2 |\mathcal{F}_k\right) \left(\sum_{l=k}^{n} b_{l-k}\right) \times \mathbb{E}\left(\sum_{l=k}^{n} |\nabla_l \Psi|^2 b_{l-k} \mid \mathcal{F}_k\right) \\
&\leq 4BM_k \sum_{l=k}^{n} \mathbb{E}\left(|\nabla_l \Psi|^2 \mid \mathcal{F}_k\right) b_{l-k}.
\end{aligned}$$

On remarque enfin que

$$\begin{aligned}
Ent(\Psi^2) &\leq \sum_{k=1}^{n} \mathbb{E}\left[M_k \log\left(\frac{M_k}{M_{k-1}}\right)\right] \\
&\leq \frac{A.C^2}{4} \sum_{k=1}^{n} \mathbb{E}\left[4.B \sum_{l=k}^{n} \mathbb{E}\left(|\nabla_l \Psi|^2 \mid \mathcal{F}_k\right) b_{l-k}\right] \\
&\leq A.C^2.B \sum_{l=1}^{n} \sum_{k=1}^{l} \mathbb{E}\left(|\nabla_l \Psi|^2\right) b_{l-k},
\end{aligned}$$

le théorème de Fubini impliquant le résultat.

Bibliographie

[1] S. Aida, Gradient estimates of harmonic functions and the asymptotics of spectral gaps on path spaces. *Interdiscip. Inform. Sci.* **2**, *no. 1, 75-84 (1996).*

[2] S. Aida, K.D. Elworthy, Differential calulus on path and loop spaces I. Logarithmic Sobolev inequalities on path spaces *C. R. Acad. Sci. Paris* **321**, *97-102 (1995).*

[3] S. Aida, T. Masuda, I. Shigekawa, Logarithmic Sobolev inequalities and exponential integrability. *J. Funct. Anal.* **126**, *no. 1, 83-101 (1994).*

[4] S. Aida, D.Stroock, Moment estimates derived from Poincaré and logarithmic Sobolev inequalities. *Math. Res. Lett.* **1**, *no. 1, 75-86 (1994).*

[5] D. Bakry, *L'hypercontractivité et son utilisation en théorie des semigroupes,* Ecole d'Eté de Probabilités de Saint-Flour 1992, LNM 1581, Springer-Verlag (1994).

[6] D. Bakry, M. Emery. Diffusions hypercontractives. Seminaire de probabilités XIX, *Lecture Notes in Math, Volume* **1123**, *177-206 (1985).*

[7] S.G. Bobkov, F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163**, *1-28 (1999).*

[8] S. Bobkov, I. Gentil, M. Ledoux. Hypercontractivity of Hamilton-Jacobi equations. *J. Math. Pure Appl.,* **80**, *no. 7, 669-696 (2001).*

[9] W.M. Boothby, *An introduction to differentiable manifolds and Riemannain geometry*, Academic Press, Orlando (1986).

[10] M. Capitaine, E.P. Hsu and M. Ledoux. Martingale representation and a simple proof of logarithmic Sobolev inequality on path spaces. *Elect. Comm. Probab.,* **2** *: paper 7 (1997).*

[11] B. Driver : A Cameron-Martin type quasi-invariance theorem for Brownan motion on a compact riemannian manifolds, *J. Funct. Anal.* **110**, *272-376 (1990).*

[12] H. Djellout, A. Guillin, Moderate deviation for Markov chains with atom. *Stochastic Processes and Their Applications,* **95**, *203-217 (2001).*

[13] H. Djellout, A. Guillin, L. Wu, Transportation cost-information inequalities and application to random dynamical systems and diffusions, *Ann. Probab.,* **32**, *2702-2732 (2004).*

[14] H. Djellout, A. Guillin, L. Wu, Moderate deviations for non-linear functionals of moving average processes, *to appear in Ann. Inst. H. Poincaré (2004).*

[15] S. Fang, Inégalité du type de Poincaré sur l'espace des chemins riemanniens, *CRAS,* **318**, *257-260 (1994).*

[16] Gentil, Inégalités de Sobolev logarithmiques et hypercontractivité en mécanique statistique et en EDP, *Thèse de Doctorat, Université de Toulouse III (2001).*

[17] L. Gross, Logarithmic Sobolev inequalities, *Amer. J. Math.,***97**, *1061-1083 (1975).*

[18] E.P. Hsu, *Stochastic Analysis on Manifolds.* Graduate Studies in Mathematics, Volume 38 (2001).

[19] E.P. Hsu, Analysis on Path and Loop Spaces. *Probability theory and applications, Princeton, NJ, 277-347 (1996)*

[20] N. Ikeda, S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North Holland, (1989).

[21] M. Ledoux. *Concentration of measure and logatithmic Sobolev inequalities.* Séminaire de Probabilités XXXIII, *Lecture Notes in Math. Springer*, 1709, 120-216 (1999).

[22] D. Nualart. *The Malliavin Calculus and related Topics*, Springer Verlag (1995).

[23] F. Otto, C. Villani Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality *J. Funct. Anal.,* **173**,*361-400 (2000).*

[24] D. Revuz , M. Yor, *Continuous Martingale and Brownian Motion, third edition*, Springer (1999).

[25] F.Y. Wang. Transportation cost inequalities on path spaces over Riemannian manifolds. *Illinois J maths* **46**, *1197-1206 (2002).*

[26] L. Wu. A deviation inequality for non reversible Markov processes. *Ann. Inst. Henri Poincaré, Probabilités et Statistiques* **36**, *4, 435-445 (2000).*

[27] L. Wu and Z.L. Zhang, Talagrand's T_2-transportation inequality w.r.t. a uniform metric for diffusions, *Acta Math. Appl. Sinica, English Series, Vol.* **20**, *No.3, 357-364 (2004).*

Chapitre 2

Logarithmic Sobolev inequalities of diffusions for the L2 metric

Travail en collaboration avec Liming Wu.

Article publié dans la revue Potential Analysis.

Abstract : Under the Bakry-Emery's Γ_2-minoration condition, we establish the logarithmic Sobolev inequality for the Brownian motion with drift in the metric L^2 instead of the usual Cameron-Martin metric. The involved constant is sharp and does not explode for large time. This inequality with respect to the L^2-metric provides us the gaussian concentration inequalities for the large time behavior of the diffusion.

Keywords : Logarithmic Sobolev inequality (LSI), concentration inequality, path space, Malliavin calculus.

AMS Subject Classification 2000: 60E15, 60H07.

2.1 Introduction

The logarithmic Sobolev inequality (LSI in short), discovered by Gross ([16], 1975) for Gaussian measures and the Wiener measure, plays a prominent role in the infinite dimensional analysis as the Sobolev inequalities do in the finite dimensional analysis. Herbst first, Aida-Masuda-Shigekawa [2] and Aida-Stroock [3] found that LSI implies the Gaussian concentration inequality for Lipschitzian functions, Bobkov-Götze [6] characterized the Gaussian concentration inequality for Lipschitzian functions by means of the so called T_1-transportation inequality, which is equivalent to the Gaussian integrability of the distance function by the recent work of H. Djellout and al. [10]. Otto and Villani [20] and Bobkov-Gentil-Ledoux [7] find that LSI implies the Talagrand T_2-transportation inequality. See Bakry [4], Ledoux [17] and Villani [21] for systematic treatment and further study.

For introducing our question for diffusions, let us first recall the LSI of Gross on the Wiener space.

Let μ be the Wiener measure on $W^d = C\big([0,T],\mathbb{R}^d\big)$ and $h \in H$ where H is the Cameron Martin space, i.e, the Hilbert space H of all absolutely continuous functions $h : [0,T] \to \mathbb{R}^d$ such that $h(0) = 0$ and

$$\|h\|_H^2 := \int_0^T |\dot{h}(t)|^2 dt < \infty.$$

The directional derivative along h of a smooth function f at $\gamma \in W^d$ is defined on W^d by :

$$D_h f(\gamma) = \lim_{\varepsilon \to 0} \frac{f(\gamma + \varepsilon h) - f(\gamma)}{\varepsilon}.$$

The linearity and continuity in h of the above expression gives the existence of an element $Df(\gamma) = (Df(\gamma,t))_{t\in[0,T]}$ in H (the Malliavin gradient) such that :

$$\forall h \in H, D_h f(\gamma) = \langle Df(\gamma), h\rangle_H = \int_0^T \dot{\widehat{Df}}(\gamma,t) \cdot \dot{h}(t)dt \tag{2.1}$$

where $\dot{\widehat{Df}}(\gamma,t) = \frac{d}{dt}Df(\gamma,t)$.

For the Wiener measure μ on W^d, Gross ([16], 1975) proved the following (see also the Ph.D of Gentil [13]):

Theorem 2.1.1. *For all bounded and smooth functions F on W^d,*

$$Ent_\mu(F^2) \leq 2C\mathbb{E}^\mu\left(\|DF\|_H^2\right) \tag{2.2}$$

where the constant $C = 1$ is sharp.

Here the entropy of $0 \leq f \in L^1(\mu)$ w.r.t. μ is defined as

$$Ent_\mu(f) = \mathbb{E}^\mu\bigg(f \log\left(\frac{f}{\mathbb{E}^\mu f}\right)\bigg).$$

When μ satisfies (2.2), we shall say that μ satisfies the LSI(C) w.r.t. the gradient in H.

This inequality w.r.t. the Cameron-Martin metric is extended to the Brownian motion over a Riemannian manifold M by Hsu [15], Aida [1], Capitaine-Hsu-Ledoux [8] under the boundedness of the Ricci curvature, and to general elliptic diffusions under the boundedness of the Bakry-Emery curvature. Particularly Aida [1] proved that the constant in the LSI does not explode for large time T under the uniform positivity of the Ricci curvature. But as explained by the second author in [23], the LSI w.r.t. the Cameron-Martin metric for diffusion (X_t) even with a constant independent of T (as in [1]) does not produce the concentration inequality of correct order in large time T for the functionals

$$F(X) := \frac{1}{\sqrt{T}}\left(\int_0^T g(X_s)ds - \mathbb{E}\int_0^T g(X_s)ds\right),\ g : M \to \mathbb{R} \text{ smooth.}$$

The reason is that the Lipschitzian coefficient of F above w.r.t. the Cameron-Martin metric is of order $\sqrt{T}$ in general, which explodes. To get useful informations for the ergodic behavior (i.e., the large time behavior) of (X_t) or more precisely for producing the Hoeffding's type Gaussian concentration inequality for functionals like $F(X)$ above for large time, as pointed out by Djellout, Guillin and the second author in [10], we should establish the LSI w.r.t. **the L^2-metric** instead of the Cameron-Martin metric, with a constant which does not explode for large time.

The famous Γ_2-minoration criterion of Bakry and Emery [5] gives the LSI for the single time law $P_T(x,\cdot) = \mathbb{P}_x(X_T \in \cdot)$ and for the unique invariant measure m of (X_t). The main purpose of this Note is to extend their LSI to the process-level $X_{[0,T]}$ w.r.t. the L^2-metric. This work is also a continuation of [10] where the Talagrand T_2-transportation inequality for diffusions w.r.t. the L^2-metric is established via the Girsanov transformation, under a weaker condition than the Bakry-Emery condition (see also [24] for T_2-inequality w.r.t. a uniform norm).

This paper is organized as follows. In the next section we first study the Brownian motion with drift, i.e., solution of the SDE

$$dX_t = dW_t + b(X_t)dt,\ X_0 = x \tag{2.3}$$

in $\mathbb{R}^d$, where the drift b satisfies the Bakry-Emery condition:

$$(\nabla b)^\sigma \le -KId,\ K \in \mathbb{R}$$

where $(\nabla b)^\sigma = ((\partial_j b_i + \partial_i b_j)/2)$ is the symmetrized gradient of the vector field b. Our method is very simple: to prove that the mapping Φ from the path W of the Brownian motion equipped with the Cameron-Martin metric to the path X of the diffusion equipped of $L^2([0,T],\mathbb{R}^d)$-norm is Lipschitzian.

The method above does not work if there is a non-constant volatility coefficient $\sigma(X_t)$ in the SDE (2.3) above. In that case, we consider a general elliptic diffusion with generator $\frac{1}{2}\Delta + V$ on a Riemannian manifold. We shall employ the method of the martingale representation formula, as developed by Fang [12] for Poincaré inequality and by Capitaine-Hsu-Ledoux [8] for log-Sobolev inequality. For the Brownian motion over a Riemannian manifold, when the Ricci curvature is bounded, we shall see in §3 how to derive the LSI quickly from the Fang's martingale representation formula. But in the general case where the Bakry-Emery's curvature is not bounded, Fang's version of martingale representation formula seems to be no longer true and that is the main difficulty. We shall use a localization technique to get a martingale representation formula for a class of test functions, and then the LSI. This is the task of §4. Finally in §5 we present several useful consequences of the LSI about the concentration of additive functionals of the diffusions for large time.

In this note we denote the maximum of u and v by $u \vee v$ and their minimum by $u \wedge v$. For $x, y \in \mathbb{R}^d$, $x \cdot y = \sum_{i=1}^d x_i y_i$ and $|x|$ denotes the Euclidian norm of x. When there is no possible confusion, we will use L^2 for $L^2([0,T];\mathbb{R}^d)$.

2.2 Brownian motion with drift on $\mathbb{R}^d$

2.2.1 Gradient ∇ in L^2 and the LSI w.r.t. ∇

We are given a standard $\mathbb{R}^d$-valued Brownian motion W_t on $\mathbb{R}^d$, a fixed initial point x in $\mathbb{R}^d$ and a continuously differentiable mapping $b : \mathbb{R}^d \to \mathbb{R}^d$. Let us consider the $\mathbb{R}^d$-valued diffusion $X_t(x)$ defined by the stochastic differential equation :

$$dX_t = dW_t + b(X_t)dt \ , \quad X_0 = x \in \mathbb{R}^d. \tag{2.4}$$

We assume the following condition on the drift b

$$(\nabla b)^\sigma \leq -K \ Id, \ \ K \in \mathbb{R} \tag{2.5}$$

where Id is the identity matrix in $\mathbb{R}^d$, $\nabla b = (\partial_j b_i)_{ij}$ is the Jacobian matrix, and A^σ denotes the symmetrization $(A + A^t)/2$ of the matrix A; and for two symmetric matrices A, B, $A \leq B$ means that $B - A$ is non-negative definite.

This condition, when $K > 0$, is exactly the Bakry-Emery's Γ_2-minoration condition for the logarithmic Sobolev inequality of the unique invariant measure m on $\mathbb{R}^d$. Details about this condition, and applications to hypercontractive diffusions can be found in [4], [5]. We are interested here in the logarithmic Sobolev inequality on the path space w.r.t. the gradient in $L^2([0,T],\mathbb{R}^d)$. More precisely, regarding $W^d = C([0,T],\mathbb{R}^d)$ as a dense subspace of $L^2([0,T],\mathbb{R}^d)$, we can introduce:

Definition 2.2.1. *For a function f on W^d, differentiable with respect to the L^2-norm, let $\nabla f(\gamma)$ be the element in $L^2([0,T],\mathbb{R}^d)$ such that*

$$D_g f(\gamma) = \langle \nabla f(\gamma) \ , \ g \rangle_{L^2([0,T],\mathbb{R}^d)} \ , \ \forall g \in H.$$

We shall write $\nabla f(\gamma) = (\nabla_t f(\gamma))_{t \in [0,T]}$.

Its relation with the gradient Df in the Cameron-Martin space H (introduced in (2.1)) is given by

Lemma 2.2.2. *For a function f on W^d, differentiable with respect to the L^2-norm, we have for every $\gamma \in W^d$, $\ddot{\widehat{D}}f(\gamma)$ is in $L^2([0,T],\mathbb{R}^d)$ in the sense of distribution and*

$$\nabla_t f(\gamma) = -\ddot{\widehat{D}}f(\gamma,t), \quad dt - a.e.$$

Proof. Such function f is differentiable w.r.t. the norm of H. Let $C_0^\infty((0,T),\mathbb{R}^d)$ be the space of all $\mathbb{R}^d$-valued C^∞-functions with compact support in $(0,T)$, which is dense in $L^2([0,T],\mathbb{R}^d)$. For every $h \in C_0^\infty((0,T),\mathbb{R}^d)$, writing $\ddot{\widehat{D}}f(\gamma,t) = \frac{d^2}{dt^2}Df(\gamma,t)$ in the sense of distribution, we have,

$$\begin{aligned}\int_0^T \nabla_s f(\gamma)\cdot h(s)ds &= \langle \nabla f(\gamma), h\rangle_{L^2} = D_h f \\ &= \int_0^T \dot{\widehat{D}}f(\gamma,s)\cdot \dot{h}(s)ds \\ &= -\int_0^T \ddot{\widehat{D}}f(\gamma,s)\cdot h(s)ds\end{aligned}$$

where the desired result follows. □

We denote by $C_b^1(W^d/L^2)$ the space of all bounded functions f on W^d, differentiable w.r.t. the $L^2([0,T],\mathbb{R}^d)$-norm, such that ∇f is also continuous and bounded from $(W^d, \|\cdot\|_{L^2})$ to $L^2([0,T],\mathbb{R}^d)$. The aim of this section is to prove

Theorem 2.2.3. *Let $\mathbb{P}_x$ be the law of the diffusion $X.(x)$ defined by the SDE (2.4). Assume (2.5). Then for all functions $f \in C_b^1(W^d/L^2)$,*

$$Ent_{\mathbb{P}_x}(f^2) \le 2C(T)\mathbb{E}^{\mathbb{P}_x}\int_0^T |\nabla_t f|^2 dt \tag{2.6}$$

where

$$C(T) := \begin{cases} \frac{1}{K^2}\left(1-\sqrt{2e^{-KT}-e^{-2KT}}\right), & \text{if } 2e^{-KT}-e^{-2KT}>0, K\neq 0\\ \frac{T^2}{2}, & \text{if } K=0 \\ \frac{1}{K^2}\left(e^{-2KT}-2e^{-KT}+1\right), & \text{if } 2e^{-KT}-e^{-2KT}\le 0, K\neq 0.\end{cases} \tag{2.7}$$

In particular if $K>0$, $C(T)\le 1/K^2$ (for all $T>0$) is non-explosive.

Note that our estimate of $C(T)$ is sharp for $T\to+\infty$ for the Ornstein-Uhlenbeck process $dX_t = dW_t - (X_t/2)dt$ by the discussion in [10, Remarks 5.7].

2.2.2 Analytic preparation

In this paragraph we recall and prove several known facts, for self-containedness. Let $M_d(\mathbb{R})$ be the space of all real $d\times d$-matrices. For $A\in M_d(\mathbb{R})$, it is clear that $\forall x\in\mathbb{R}^d$, $\langle Ax\ ,\ x\rangle=\langle A^\sigma x\ ,\ x\rangle$. We first give a Gronwall Lemma of matrix type.

Lemma 2.2.4. *If $J(t)\in M_d(\mathbb{R})$ satisfies the equation*

$$\dot{J}(t)=A(t)J(t)\quad \textit{with}\quad J(0)=Id,$$

and if $A^\sigma\le BId$ for a real number B in the order of the nonnegative definiteness, then we have

$$|J(t)y|\le e^{Bt}|y|,\forall y\in\mathbb{R}^d,\ \forall t\ge 0.$$

Proof. Indeed we have :

$$\begin{aligned}\frac{d}{dt}\big(e^{-2Bt}|J(t)x|^2\big) &= e^{-2Bt}\big(-2B|J(t)x|^2+2\langle\dot{J}(t)x,J(t)x\rangle\big)\\ &= e^{-2Bt}\big(-2B|J(t)x|^2+2\langle A(t)J(t)x,J(t)x\rangle\big)\\ &= e^{-2Bt}\big(-2B|J(t)x|^2+2\langle A^\sigma(t)J(t)x,J(t)x\rangle\big)\\ &= -2e^{-2Bt}\langle\left(BId-A^\sigma(t)\right)J(t)x,J(t)x\rangle\\ &\le 0\end{aligned}$$

so,

$$|J(t)x|\le e^{Bt}|x|.$$

□

We now give a result to control the norm of a bounded self-adjoint operator from L^2 to L^2.

Lemma 2.2.5. *Let Γ be a bounded self-adjoint operator from $L^2(E,\mu)$ to $L^2(E,\mu)$ where μ is σ-finite. Assume that*

$$\|\Gamma\|_{1,1}:=\sup_{0\neq f\in L^1\bigcap L^2}\frac{\|\Gamma f\|_{L^1(\mu)}}{\|f\|_{L^1(\mu)}}<+\infty.$$

Then $\|\Gamma\|_{2,2}$(the norm in L^2) $\le\|\Gamma\|_{1,1}$.

Proof. For any $f\in L^1\bigcap L^\infty\subset L^2$, we have by the symmetry of Γ,

$$\begin{aligned}\|\Gamma f\|_\infty &= \sup_{g\in L^1\bigcap L^\infty,\|g\|_1\le 1}\langle\Gamma f,g\rangle\\ &= \sup_{g\in L^1\bigcap L^\infty,\|g\|_1\le 1}\langle f,\Gamma g\rangle\\ &\le\|\Gamma\|_{1,1}\|f\|_\infty,\end{aligned}$$

i.e., $\|\Gamma\|_{\infty,\infty}\le\|\Gamma\|_{1,1}$. Hence by the Riesz-Thorin's interpolation theorem, $\|\Gamma\|_{2,2}\le\|\Gamma\|_{1,1}$. □

2.2.3 Proof of Theorem 2.2.3

Outline of the proof

The SDE (2.4) has a unique strong solution (the reader is referred to [18], p 194)

$$X = \Phi(W),$$

where $\Phi : W^d \to W^d$. Indeed Φ is continuous on W^d and for each $\gamma \in W^d$, $\eta = \Phi(\gamma)$ is the unique solution of the ordinary differential equation

$$\eta_t = x + \gamma_t + \int_0^t b(\eta_s)ds.$$

We start with the LSI (2.2) of Gross. If $f \circ \Phi \in \mathbb{D}_2(D)$ (the domain of D in $L^2(W^d, \mu)$ as defined in the Malliavin calculus), we have

$$Ent_{\mathbb{P}_x} f^2 = Ent_\mu[(f \circ \Phi)^2] \leq 2\mathbb{E}^\mu \|D(f \circ \Phi)\|_H^2. \tag{2.8}$$

Now for every $h \in H$, we shall prove that

$$A(\gamma)h := \lim_{\varepsilon \to 0} \frac{\Phi(\gamma + \varepsilon h) - \Phi(\gamma)}{\varepsilon} \tag{2.9}$$

holds in the sup-norm of W^d (then in that of L^2). Thus

$$\Phi(\gamma + \varepsilon h) = \Phi(\gamma) + \varepsilon \frac{d\Phi(\gamma + \varepsilon h)}{d\varepsilon}|_{\varepsilon=0} + \circ(\varepsilon) = \Phi(\gamma) + \varepsilon A(\gamma)h + \circ(\varepsilon)$$

and consequently

$$\begin{aligned} D_h(f \circ \Phi)(\gamma) &= \lim_{\varepsilon \to 0} \frac{f\big(\Phi(\gamma + \varepsilon h)\big) - f\big(\Phi(\gamma)\big)}{\varepsilon} \\ &= \lim_{\varepsilon \to 0} \frac{f\big(\Phi(\gamma) + \varepsilon A(\gamma)h + \circ(\varepsilon)\big) - f\big(\Phi(\gamma)\big)}{\varepsilon} \\ &= (D_{A(\gamma)h} f)\big(\Phi(\gamma)\big). \end{aligned}$$

If we can prove that

$$\|A(\gamma)h\|_{L^2} \leq \sqrt{C(T)}\|h\|_H \tag{2.10}$$

where $C(T)$ is given in the theorem, we shall get

$$\begin{aligned} \|D(f \circ \Phi)(\gamma)\|_H &= \sup_{\|h\|_H \leq 1} |D_h(f \circ \Phi)(\gamma)| \\ &= \sup_{h \in H\ ,\ \|h\|_H \leq 1} |(D_{A(\gamma)h} f)\big(\Phi(\gamma)\big)| \\ &= \sup_{h \in H\ ,\ \|h\|_H \leq 1} \langle (\nabla f)\big(\Phi(\gamma)\big)\ ,\ A(\gamma)h\rangle_{L^2} \end{aligned}$$

$$\begin{aligned}
&\leq \sup_{h\in H\ ,\ \|h\|_H\leq 1} \|(\nabla f)\big(\Phi(\gamma)\big)\|_{L^2}\|A(\gamma)h\|_{L^2}\\
&\leq \sqrt{C(T)}\|(\nabla f)\big(\Phi(\gamma)\big)\|_{L^2}.
\end{aligned}$$

Thus by (2.8), we obtain

$$\begin{aligned}
Ent_{\mathbb{P}_x} f^2 &\leq 2\mathbb{E}^\mu\|D(f\circ\Phi)\|_H^2\\
&\leq 2C(T)\mathbb{E}^\mu\|(\nabla f)\circ\Phi\big)\|_{L^2}^2\\
&= 2C(T)\mathbb{E}^{\mathbb{P}_x}\int_0^T |\nabla_t f|^2 dt
\end{aligned}$$

the desired inequality.

Proof of Theorem 2.2.3

To make the above proof rigorous, it remains to verify three points: (2.9), (2.10) and $f\circ\Phi\in\mathbb{D}_2(D)$. The last point is a consequence of (2.9), (2.10), because (2.9), (2.10) imply

$$|D_h(f\circ\Phi)(\gamma)| \leq \sqrt{C(T)}\|h\|_H\|(\nabla f)(\Phi(\gamma))\|_{L^2}$$

for all $h\in H$ (this is a well known fact in the Malliavin calculus, cf. Nualart [19]). (2.9) is an elementary fact in the ODE theory, so omitted. We now prove (2.10).

Since $\eta^\epsilon = \Phi(\gamma+\varepsilon h)$ verifies

$$\eta_t^\epsilon = x + \gamma_t + \varepsilon h(t) + \int_0^t b\big(\eta_s^\varepsilon\big)ds$$

then, $g := A(\gamma)h = \frac{d}{d\varepsilon}\eta^\varepsilon$ satisfies the linear differential equation :

$$g(t) = h(t) + \int_0^t \nabla b\big(\eta_s\big)g(s)ds$$

where $\eta = \Phi(\gamma)$. This equation has a unique solution and it can be given explicitly by

$$g(t) = \int_0^t A_t{A_s}^{-1}\dot{h}(s)ds$$

where A_t satisfies :

$$A_0 = Id\ ,\quad \frac{d}{dt}A_t = \nabla b(\eta_t)A_t.$$

Under the condition (2.5), we obtain using Lemma 2.2.4 :

$$|A_t{A_s}^{-1}x| \leq e^{-K(t-s)}|x|,\ \forall 0\leq s\leq t.$$

We state now the desired inequality (2.10) as

Lemma 2.2.6. *We have*

$$\|A(\gamma)h\|_{L^2}^2 = \|g\|_{L^2}^2 \leq C(T)\|h\|_H^2$$

where $C(T)$ is given in Theorem 2.2.3.

Proof. From the following expression $g(t) = \int_0^t A_t A_s^{-1}\dot{h}(s)ds$, we can compute :

$$\begin{aligned} \|g\|_{L^2}^2 &= \int_0^T \left|\int_0^t A_t A_s^{-1}\dot{h}(s)ds\right|^2 dt \\ &\leq \int_0^T \left(\int_0^t e^{-K(t-s)}|\dot{h}(s)|ds\right)^2 dt \end{aligned}$$

Now, we have to control the last term. A Cauchy-Schwarz control is easy, but the constant so obtained explodes. In order to avoid this explosion, let us write

$$\begin{aligned} &\int_0^T \left(\int_0^t e^{-K(t-s)}|\dot{h}(s)|ds\right)^2 dt \\ = &\int_0^T \left(\int_0^t e^{-K(t-u)}|\dot{h}(u)|du\right)\left(\int_0^t e^{-K(t-v)}|\dot{h}(v)|dv\right) dt \\ = &\int_0^T \left(\iint_{0\leq u,v\leq t} e^{-K(t-u)}|\dot{h}(u)|e^{-K(t-v)}|\dot{h}(v)|du\ dv\right) dt \\ = &\iint_{0\leq u,v\leq T} \left(\int_{u\vee v}^T e^{-2Kt+K(u+v)}dt\right)|\dot{h}(u)||\dot{h}(v)|du\ dv. \end{aligned}$$

Since $2(u\vee v)-(u+v) = |u-v|$, we have if $K \neq 0$,

$$\begin{aligned} \Gamma(u,v) := \int_{u\vee v}^T e^{-2Kt+K(u+v)}dt &= \frac{e^{K(u+v)}}{2K}\left(e^{-2K(u\vee v)} - e^{-2KT}\right) \\ &= \frac{e^{-K|u-v|} - e^{-K(2T-u-v)}}{2K} \end{aligned}$$

and if $K = 0$,

$$\Gamma(u,v) := \int_{u\vee v}^T e^{-2Kt+K(u+v)}dt = (T - u\vee v).$$

Then we obtain :

$$\|g\|_{L^2}^2 \leq \int_0^T\int_0^T \Gamma(u,v)|\dot{h}(u)||\dot{h}(v)|dudv. \qquad (2.11)$$

Define an operator $\Gamma : L^2(0,T) \to L^2(0,T)$ by :

$$\Gamma f(v) = \int_0^T \Gamma(u,v)f(u)du.$$

It is a self-adjoint operator on $L^2(0,T)$. Then the term to control becomes :

$$\langle \Gamma |\dot{h}| \ , \ |\dot{h}| \rangle_{L^2}.$$

But it is clear that (since $\Gamma(u,v) \geq 0$):

$$\begin{aligned} \|\Gamma f\|_{L^1} &= \int_0^T \left| \int_0^T \Gamma(u,v) f(u) du \right| dv \\ &\leq \left(\sup_{v\in[0,T]} \int_0^T \Gamma(u,v) du \right) \int_0^T |f(v)| dv. \end{aligned}$$

Thus Lemma 2.2.5 yields $\|\Gamma\|_{L^2} \leq A(T)$ where

$$A(T) = \sup_{v\in[0,T]} \int_0^T \Gamma(u,v) du.$$

Finally by an elementary calculus, we get $A(T) = C(T)$ given in Theorem 2.2.3.

□

Remarks 2.2.7. In the proof above, we see that $(\gamma, h) \to A(\gamma)h = g$ is continuous from $W^d \times H$ to $L^2([0,T], \mathbb{R}^d)$. Thus for all $f \in C_b^1(W^d/L^2)$, $f \circ \Phi$ is continuously differentiable on W^d w.r.t. the norm of H, and we have shown

$$\|D(f \circ \Phi)(\gamma)\|_H \leq \sqrt{C(T)} \|\nabla f(\Phi(\gamma))\|_H.$$

Now let $Q \in M_d(\mathbb{R})$ and F a continuously differentiable function on W^d w.r.t. the norm of H, applying the Gross theorem to $F(Q\gamma)$, we get

$$Ent(F^2(QW.)) \leq 2\lambda_{max}(Q^tQ)\mathbb{E} \int_0^T |\dot{\widehat{DF}}|^2(QW., t) dt$$

where $\lambda_{max}(Q^tQ)$ is the maximal eigenvalue of Q^tQ. Thus by the estimate above, we see that

$$dX_t = b(X_t)dt + QdW_t, \ X_0 = x$$

satisfies : $\forall f \in C_b^1(W^d/L^2)$,

$$Ent(f^2(X)) \leq 2C(T)\lambda_{max}(Q^tQ)\mathbb{E} \int_0^T |\nabla_t f|^2(X) dt.$$

In other words our approach here is also well adapted for eventually degenerate noise QdW_t.

2.3 Brownian motion on manifold

When one studies the diffusion $dX_t = \sigma(X_t)dW_t + b(X_t)dt$ with a non-constant volatility coefficient σ, the application Φ mapping the path of the BM to the path of X does not satisfy : $\|\nabla_h \Phi\|_{L^2} \leq \sqrt{C(T)}\|h\|_H$. So the approach in the previous section loses its pertinence. Instead, we shall explore another method, based on the martingale representation formula, as developed in [12] and [8]. We shall study only the elliptic diffusion on a Riemannian manifold generated by $\frac{1}{2}\Delta + V$.

To illustrate clearly the idea, we begin with the BM over a Riemannian manifold with bounded Ricci curvature. In that case, using the martingale representation formula, Fang [12] obtained the Poincaré inequality, Hsu [14, 15] and Capitaine-Hsu-Ledoux [8] obtained the logarithmic Sobolev inequality on the path space, both w.r.t. the Cameron-Martin metric. The constant of the LSI in those works explodes for large time t. Aida [1] obtained the LSI w.r.t. the Cameron-Martin metric with a bounded constant for large time T once the Ricci curvature is bounded from below by a positive constant. But as explained in the Introduction the LSI w.r.t. the Cameron-Martin metric even with a non-explosive constant does not produce concentration inequality of correct order (in large time T) for functionals such as $F(X) = \int_0^T g(X_s)ds$.

2.3.1 Martingale representation and gradient w.r.t. L^2-norm

We follow the exposition of [8]. Let M be a complete and connected manifold of dimension d equipped with the Levi-Civita connection ∇. Denote by Δ the Laplace-Beltrami operator on M, $O(M)$ the bundle of orthonormal frames, and $\Pi : O(M) \to M$ the canonical projection. Fix a point $x_0 \in M$. Each frame $u \in O(M)$ is a linear isometry $u : \mathbb{R}^d \to T_{\Pi(u)}(M)$. We assume throughout this section the boundedness of the Ricci curvature

$$\sup_{u \in O(M)} \|Ric_u\| < +\infty, \tag{2.12}$$

and denote also Ric_u its scalarization. Let $W_{x_0}(M)$ be the space of continuous paths from $[0,T]$ to M starting at x_0. Let U be the horizontal Brownian motion. We recall that U is solution of the Stratonovich SDE :

$$dU_t = \sum_{i=1}^{d} H_i(U_t) \circ dW_t^i \ , \ U_0 = u_0 \in O(M) \tag{2.13}$$

where W_t is a standard Brownian motion on $\mathbb{R}^d$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $H_i, 1 \leq i \leq n$ are the canonical horizontal vector fields on $O(M)$. Details about this construction can be found in [14] and [15]. Then $X = \Pi(U)$ is the M-valued Brownian motion starting at x_0, whose law μ_M is the Wiener measure on $W_{x_0}(M)$. Hence, for a Brownian path X, and h in the Cameron Martin space, $U_t h_t$ is a tangent vector at X_t. It determines a vector field D_h by $D_h X_t = U_t h_t$.

For smooth cylindrical functions f on $W_{x_0}(M)$, that is $f(\gamma) = f(\gamma_{t_1}, \ldots, \gamma_{t_n})$ with $\gamma \in W_{x_0}(M)$, the Malliavin derivative operator is defined by :

$$Df(X,t) = \sum_{i=1}^{n} U_{t_i}^{-1} \frac{df}{dx_i}(X_{t_1}, \ldots, X_{t_n}) t_i \wedge t.$$

Notice that from the path of the BM X over the Riemannian manifold and $U_0 = u_0$, one can reconstruct (U_t) and (W_t) a.s. (w.r.t. the law of X). Thus the above definition is intrinsic (cf. [14]).

Let $(D, \mathbb{D}_2(D))$ be its closure in $L^2(\mu_M)$. For $h \in H$, $D_h f(\gamma)$ and $Df(\gamma) \in H$ are such that :

$$\forall h \in H,\ D_h f(\gamma) = \langle Df(\gamma)\ ,\ h \rangle_H = \int_0^T \dot{\widehat{D}} f(\gamma, s) \cdot \dot{h}(s) ds.$$

Let $F : W_{x_0}(M) \to \mathbb{R}$ be a real function in the domain $\mathbb{D}_2(D)$ of D. Under the boundedness condition (2.12), the Fang's version of Clark-Ocone-Haussmann formula for the Brownian motion on a manifold is read as

$$F(X) = \mathbb{E}\big(F(X)\big) + \int_0^T \langle H_t\ , dW_t \rangle \tag{2.14}$$

where

$$H_t = \mathbb{E}^{\mu_M}\left(\dot{\widehat{D}} F(t) - \frac{1}{2} A_t^* \int_t^T (A_s^*)^{-1} Ric_{U_s} \dot{\widehat{D}} F(s) ds | \mathcal{B}_t \right). \tag{2.15}$$

In this expression, $(\mathcal{B}_t)$ is the filtration generated by (W_t), $(A_t)_{0 \le t \le T}$ is a matrix valued process defined by :

$$\frac{dA_t}{dt} - \frac{1}{2} A_t Ric_{U_t} = 0 \ \ ,\ \ A_0 = Id. \tag{2.16}$$

The above formula relies on the Bismut-Driver integration by parts formula : for any $h \in H$, the adjoint D_h^* of D_h in $L^2(\mu_M)$ is given by

$$D_h^* = -D_h + \int_0^T \langle \dot{h}_t + \frac{1}{2} Ric_{U_t} h_t, dW_t \rangle. \tag{2.17}$$

According to Lemma 2.2.2, we introduce our gradient w.r.t. the L^2-metric.

Definition 2.3.1. *For $F \in \mathbb{D}_2(D)$ such that $\ddot{\widehat{DF}} \in L^2([0,T],\mathbb{R}^d)$ we define the gradient of F w.r.t. L^2-norm by*

$$\nabla_t F(\gamma) = -\ddot{\widehat{DF}}(\gamma,t) \quad dt - a.e.$$

In other words, $\nabla F(\gamma) \in L^2([0,T],\mathbb{R}^d)$ is determined by

$$\langle \nabla F(\gamma)\ ,\ h\rangle_{L^2} \ = \ D_h F(\gamma)\ ,\ \forall h \in C_0^\infty(]0,T[,\mathbb{R}^d).$$

We assume now

$$\forall u \in O(M)\ ,\quad \frac{1}{2}Ric_u \geq KId,\ \ K \in \mathbb{R}. \tag{2.18}$$

The constant of our LSI will depend only on this lower bound on the Ricci curvature.

Lemma 2.3.2. *Let A solution of (2.16). We have*

$$\forall s \geq t\ ,\quad \|A_t^* A_s^{-1^*}\| \leq \exp\big(-K(s-t)\big).$$

Proof. With a derivation of $A_t A_t^{-1} = Id$, we obtain

$$\frac{d}{ds}(A_s^{-1}A_t) = -\frac{1}{2}Ric_{U_s}A_s^{-1}A_t,$$

which implies the desired result using Lemma 2.2.4 and (2.18). □

The key observation is:

Lemma 2.3.3. *For $F \in \mathbb{D}_2(D)$ such that $\mu_M - a.s$, $\ddot{\widehat{DF}}(\gamma) \in L^2([0,T],\mathbb{R}^d)$ (then $\dot{\widehat{DF}}(\gamma)$ can and will be chosen to be absolutely continuous) and $\dot{\widehat{DF}}(\gamma,T) = 0$, in the representation formula (2.14), we have*

$$H_t = \mathbb{E}^{\mu_M}\left(\int_t^T A_t^* A_s^{*-1}\nabla_s F ds|\mathcal{B}_t\right).$$

Proof. Taking the adjoint in the proof of Lemma 2.3.2 gives

$$\frac{d}{ds}(A_t^*A_s^{*-1}) = -\frac{1}{2}A_t^*A_s^{*-1}Ric_{U_s}.$$

By integration by parts in the expression (2.15) of H_t, we have $\mu_M - a.s.$,

$$\begin{aligned}
-\frac{1}{2}A_t^*\int_t^T (A_s^*)^{-1}Ric_{U_s}\dot{\widehat{DF}}(s)ds \ &= \ \int_t^T \frac{d}{ds}(A_t^*A_s^{*-1})\dot{\widehat{DF}}(s)ds \\
&= \ \Big[A_t^*A_s^{*-1}\dot{\widehat{DF}}(s)\Big]_{s=t}^{s=T} - \int_t^T A_t^*A_s^{*-1}\ddot{\widehat{DF}}(s)ds \\
&= \ -\dot{\widehat{DF}}(t) + \int_t^T A_t^*A_s^{*-1}\nabla_s F ds.
\end{aligned}$$

Then the desired result follows by (2.15). □

2.3.2 Poincaré inequality for the L^2 norm

Now, we follow Fang [12]. For each $F \in \mathbb{D}_2(D)$ such that $\mu_M - a.s$, $\widehat{\ddot{D}F}(\gamma) \in L^2([0,T],\mathbb{R}^d)$ and $\widehat{\dot{D}F}(\gamma,T) = 0$, we have by Lemma 2.3.3,

$$\begin{aligned} Var_{\mu_M}(F) &= \mathbb{E}\left(F(X) - \mathbb{E}F(X)\right)^2 = \mathbb{E}^{\mu_M}\int_0^T |H_t|^2 dt \\ &= \mathbb{E}\int_0^T \left|\mathbb{E}\left[\int_t^T A_t^* A_s^{*-1} \nabla_s F(X) ds | \mathcal{B}_t\right]\right|^2 dt \\ &\leq \mathbb{E}\int_0^T \left[\int_t^T \|A_t^* A_s^{*-1}\| \, |(\nabla_s F)(X)| ds\right]^2 dt \\ &\leq \mathbb{E}^{\mu_M}\int_0^T \left[\int_t^T e^{-K(s-t)} |\nabla_s F| ds\right]^2 dt \end{aligned}$$

where the last inequality is given by Lemma 2.3.2. We write the last quantity as (by Fubini)

$$\begin{aligned} &\mathbb{E}^{\mu_M}\int_0^T \int_t^T e^{-K(u-t)} |\nabla_u F| \int_t^T e^{-K(v-t)} |\nabla_v F| du dv dt \\ &\quad = \mathbb{E}^{\mu_M}\int_0^T \int_0^T du dv |\nabla_u F| \cdot |\nabla_v F| \int_0^{u\wedge v} e^{-K(u-t)-K(v-t)} dt. \end{aligned}$$

Let $\Gamma(u,v) := \int_0^{u\wedge v} e^{-K(u-t)-K(v-t)} dt$ and $\Gamma f(v) := \int_0^T \Gamma(u,v) f(u) du$. We have as in Lemma 2.2.6

$$\|\Gamma\|_{2,2} \leq \|\Gamma\|_{1,1} = \sup_{u\in[0,T]} \int_0^T \Gamma(u,v) dv \leq C(T)$$

where $C(T)$ is given in Theorem 2.2.3. Thus we get,

$$Var_{\mu_M}(F) \leq \|\Gamma\|_{2,2} \mathbb{E}^{\mu_M}\int_0^T |\nabla_t F|^2 dt \leq C(T)\mathbb{E}^{\mu_M}\int_0^T |\nabla_t F|^2 dt,$$

i.e., we have shown

Theorem 2.3.4. *Assume (2.12) and (2.18). For each $F \in \mathbb{D}_2(D)$ such that $\mu_M - a.s$, $\widehat{\ddot{D}F}(\gamma) \in L^2([0,T],\mathbb{R}^d)$ and $\widehat{\dot{D}F}(\gamma,T) = 0$,*

$$Var_{\mu_M}(F) \leq C(T)\mathbb{E}^{\mu_M}\int_0^T |\nabla_t F|^2 dt$$

where $C(T)$ is given in Theorem 2.2.3.

2.3.3 The logarithmic Sobolev inequality for L^2-norm

Theorem 2.3.5. *Assume (2.12) and (2.18). For each $F \in \mathbb{D}_2(D)$ such that μ_M- a.s, $\ddot{\widehat{D}}F(\gamma) \in L^2([0,T],\mathbb{R}^d)$ and $\dot{\widehat{D}}F(\gamma,T)=0$,*

$$Ent_{\mu_M}\big(F^2\big) \le 2C(T)\mathbb{E}^{\mu_M}\int_0^T |\nabla_t F|^2 dt$$

where $C(T)$ is given in Theorem 2.2.3.

Proof. We follow the ingenious method of Capitaine-Hsu-Ledoux [8]. Assume at first

$$\frac{1}{\varepsilon} \ge F \ge \varepsilon > 0 \tag{2.19}$$

for some $\varepsilon \in (0,1)$. Consider the continuous martingale $M_s = \mathbb{E}\left[F^2(X) \mid \mathcal{B}_s\right]$, such that $M_0 = \mathbb{E}^{\mu_M}F^2$ and $M_T = F^2(X)$. Let

$$H_t = \mathbb{E}^{\mu_M}\left(\int_t^T A_t^* A_s^{*-1}\nabla_s(F^2)ds|\mathcal{B}_t\right).$$

Using the martingale representation in Lemma 2.3.3 and applying the Itô formula, we get after taking expectation

$$\mathbb{E}\left(M_T\log M_T\right) - \mathbb{E}\left(M_0\log M_0\right) = \mathbb{E}\frac{1}{2}\int_0^T \frac{d\langle M\rangle_s}{M_s} = \mathbb{E}\frac{1}{2}\int_0^T \frac{|H_s|^2}{M_s}ds.$$

Since $\nabla_s(F^2) = 2F\nabla_s F$, $ds \otimes \mu_M - a.e.$, by the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
Ent_{\mu_M}(F^2) &= \frac{1}{2}\mathbb{E}\int_0^T \frac{\left|\mathbb{E}\left(\int_t^T A_t^*A_s^{*-1}\nabla_s(F^2)(X)ds \mid \mathcal{B}_t\right)\right|^2}{\mathbb{E}\left[F^2(X) \mid \mathcal{B}_t\right]}dt \\
&\le 2\mathbb{E}\int_0^T \mathbb{E}\left[\left|\int_t^T A_t^*A_s^{*-1}\nabla_s F(X)ds\right|^2 \mid \mathcal{B}_t\right]dt \\
&\le 2\mathbb{E}^{\mu_M}\int_0^T\left(\int_t^T \|A_t^*A_s^{*-1}\|\cdot|\nabla_s F|ds\right)^2 dt.
\end{aligned}$$

We can conclude by the proof of the Poincaré inequality.

We now remove the restriction (2.19). For general F, let

$$F_{\varepsilon,L} := f_{\varepsilon,L}(F), \text{ where } f_{\varepsilon,L} := L\arctan\frac{\sqrt{\varepsilon + x^2}}{L}.$$

It is easy to see that $F_{\varepsilon,L} \in \mathbb{D}_2(D)$ and

$$\dot{\widehat{D}}F_{\varepsilon,L}(t) = f'_{\varepsilon,L}(F)\dot{\widehat{D}}F(t),\ \nabla_t F_{\varepsilon,L} = f'_{\varepsilon,L}(F)\nabla_t F.$$

Then $F_{\varepsilon,L}$ satisfies again the assumption of the theorem and especially (2.19), so the LSI is true for $F_{\varepsilon,L}$. Moreover by the expression above we also have

$$\mathbb{E}^{\mu_M}\int_0^T |\nabla_t F_{\varepsilon,L}|^2 dt \le \mathbb{E}^{\mu_M}\int_0^T |\nabla_t F|^2 dt.$$

By letting $L \to +\infty$ first $F_{\varepsilon,L}^2 \uparrow \varepsilon + F^2$, and $\varepsilon \to 0+$ next, we have $Ent_{\mathbb{P}}\left(F_{\varepsilon,L}^2(X)\right) \to Ent_{\mathbb{P}}\left(F^2(X)\right)$, so the desired LSI for F follows. $\square$

Notice that the smooth cylindrical functions $f(\gamma_{t_1}, \cdots, \gamma_{t_n})$ does not satisfy the assumptions of Theorems 2.3.4 and 2.3.5 (this is quite natural as for t fixed, γ_t has no meaning in L^2). However $F(\gamma) := \int_0^T f(t,\gamma_t)dt$ with bounded and smooth $f : [0,T] \times M \to \mathbb{R}$ satisfies the conditions in Theorem 2.3.4 and Theorem 2.3.5, because it is easy to prove that $F \in \mathbb{D}_2(D)$ and

$$\widehat{\dot{D}F}(t,X) = \int_t^T U_s^{-1}\nabla_x f(s,X_s)ds,\ \nabla_t F(X) = U_t^{-1}\nabla_x f(t,X_t)$$

where $\nabla_x f(t,x)$ is the gradient on the space variable $x \in M$.

Before moving to our true object, let us make some comments on the condition (2.12) of the boundedness of the Ricci curvature, whereas the final results in this section seem not to depend on it "in appearance".

Remarks 2.3.6. *If the Ricci curvature is only lower bounded but not bounded, we have several technical questions related with the stochastic calculus used in this section:*

(a) The Bismut-Driver integration by parts formula (2.17) has a question with the unbounded term $Ric_{U_t}h_t$.

(b) Without the integration by parts formula (2.17), we have neither the Fang's martingale representation formula (which also contains an unbounded term), nor the closability of D in $L^2(\mu_M)$ (more exactly we do not know).

The last point is particularly annoying, since even $\mathbb{D}_2(D)$ is no longer well defined.

Now let us see how to bypass those delicate difficulties for general diffusions generated by $\frac{1}{2}\Delta + V$, whose Bakry-Emery's curvature is only assumed to be lower bounded.

2.4 Brownian motion with drift over a Riemannian manifold

In this section we extend the preceding logarithmic Sobolev inequality to the law of a diffusion process (X_t) with generator $L = \frac{1}{2}\Delta + V$ over a complete connected Riemannian manifold M with lower bounded Bakry-Emery curvature, i.e.,

$$\frac{1}{2} Ric_u - (\overline{\nabla V})^\sigma \geq KId, \ K \in \mathbb{R}. \tag{2.20}$$

Here V is a smooth vector field and $(\nabla V)^\sigma$ is the symmetrized gradient. Recall that Wang [22] has obtained the Talagrand "T_2"-transportation inequality.

We assume that the BM is non-explosive. In that situation, the diffusion with generator $L = \frac{1}{2}\Delta + V$ is non-explosive (under (2.20)) and the law μ^V of the diffusion is absolutely continuous w.r.t. the law of the BM μ_M.

2.4.1 A space of test-functions $\mathcal{D}$

We introduce now a space $\mathcal{D}$ of test-functions which is rich enough for applications in the concentration phenomena of the diffusion for large time. It allows us to circumvent the technical problems related to the fact mentioned in Remarks 2.3.6 that $\mathbb{D}_2(D)$ is ill-posed.

Definition 2.4.1. *Let*

$$\mathcal{D} = \left\{ \ F(\gamma) \ = \ \Phi\left(\int_0^T f_1(s,\gamma_s)ds \ , \ \ldots \ , \ \int_0^T f_m(s,\gamma_s)ds \right) \ \right\}$$

where $m \in \mathbb{N}^*$, $\Phi \in \mathcal{C}_b^\infty(\mathbb{R}^m, \mathbb{R})$ and $f_i \in \mathcal{C}_0^\infty([0,T] \times M)$.

We define the gradient ∇ w.r.t. L^2-norm by

$$\nabla_t \int_0^T f(s,\gamma_s)ds = U_t^{-1} \nabla_x f(t,\gamma_t)$$

and the fact that it obeys the rule of composition:

$$\nabla_t \Phi\left(\int_0^T f_1(s,\gamma_s)ds \ , \ \ldots \ , \ \int_0^T f_m(s,\gamma_s)ds \right) := \sum_i \partial_i \Phi(\cdots) U_t^{-1} \nabla_x f_i(t,\gamma_t).$$

Here $\nabla_x f(t,x)$ is the gradient w.r.t. the space variable $x \in M$.

Notice that the stochastic parallel transport $U_t = U_t(\gamma)$ is $\mu_M - a.s.$-well defined, then $\mu^V - a.s.$ well defined.

2.4.2 Martingale representation for test-functions in $\mathcal{D}$

For $u \in O(M)$, the bundle of orthonormal frames, let $\overline{V}(u) = u^{-1}V(\Pi(u)) \in \mathbb{R}^d$, and let $\overline{\nabla V}(u) = u^{-1}\nabla V(\Pi(u)) \in \mathbb{R}^d \otimes \mathbb{R}^d$. They are the scalarization of the tensors V and ∇V respectively.

Recall that our diffusion (X_t) with $X_0 = x_0$ can be constructed as follows: $X_t = \Pi(U_t)$ where (U_t) is the $O(M)$-valued diffusion, solution of

$$dU_t = \sum_{i=1}^{n} H_i(U_t) \circ dW_t^i + \tilde{V}(U_t)dt \tag{2.21}$$

where $\tilde{V}$ is the horizontal lift of V, defined on $(\Omega, (\mathcal{B}_t), \mathbb{P})$ and $(\mathcal{B}_t)$ is the filtration of the $\mathbb{R}^d$-valued Brownian motion (W_t). Notice that solution U_t of (2.21) coincides in law with $U_t(X)$ where this last U_t is used in the definition of $\nabla_t F$.

Lemma 2.4.2. *(due to [8]) Assume that Ric_u and ∇V are both bounded. Then for any smooth cylindrical functional $F = F(\gamma_{t_1}, \cdots, \gamma_{t_n})$ on $W_{x_0}(M)$,*

$$F(X) = \mathbb{E}F(X) + \int_0^T \langle H_t, dW_t\rangle, \tag{2.22}$$

where

$$\begin{aligned} H_t &= \mathbb{E}^{\mu_V}\left(\dot{\widehat{D}F}(t) - \int_t^T A_t^*(A_s^*)^{-1}M_s^*\dot{\widehat{D}F}(s)ds|\mathcal{B}_t\right) \\ M_t &= \frac{1}{2}Ric_{U_t} - \overline{\nabla V}(U_t) \\ A_0 &= I,\ \frac{dA_t}{dt} = A_tM_t. \end{aligned} \tag{2.23}$$

In the actual bounded case, it is easy to see that $F \in \mathcal{D}$ satisfies again (2.22). By integration by parts, we obtain as in Lemma 2.3.3,

Lemma 2.4.3. *Assume that Ric_u and ∇V are both bounded. Then for any $F \in \mathcal{D}$,*

$$F(X) = \mathbb{E}F(X) + \int_0^T \langle \mathbb{E}\left(\int_t^T A_t^*(A_s^*)^{-1}\nabla_s F(X)|\mathcal{B}_t\right), dW_t\rangle,$$

where (A_t) is given in Lemma 2.4.2.

Now we remove the boundedness condition.

Theorem 2.4.4. *Assume (2.20). Then the conclusion of Lemma 2.4.3 remains true.*

As corollaries of the above martingale representation, we obtain by repeating the arguments in §3:

Theorem 2.4.5. *For any $F \in \mathcal{D}$,*

$$Var_{\mu^V}(F) \le C(T)\mathbb{E}^{\mu_V} \int_0^T |\nabla_t F|^2 dt, \tag{2.24}$$

$$Ent_{\mu^V}(F^2) \le 2C(T)\mathbb{E}^{\mu_V} \int_0^T |\nabla_t F|^2 dt, \tag{2.25}$$

where $C(T)$ is given in Theorem 2.2.3.

2.4.3 Proof of Theorem 2.4.4: technique of localization

If M is compact, this result is contained in Lemma 2.4.3. Assume below that M is non-compact.

Step 1: localization. Let $F = \Phi\left(\int_0^T f_1(s,\gamma_s)ds \ , \ \dots \ , \ \int_0^T f_m(s,\gamma_s)ds\right) \in \mathcal{D}$ and fix some compact $K \subset M$ such that the supports of $f_i,\ i = 1, \cdots, m$ are all contained in $[0,T] \times K$. Take next a sequence of compact subsets $(K_n)_{n\ge1}$ of M such that

$$K \subset K_1^o(\text{ interior of } K_1),\ K_n \subset K_{n+1}^o,\ \bigcup_n K_n = M.$$

For each $n \ge 1$, choose

(1) a complete connected Riemannian manifold $M_n \supset K_n$ such that the Riemannian metric of M_n restricted on K_n coincides with the original one of M and the Ricci curvature of M_n is globally bounded;

(2) a smooth vector field V_n on M_n such that $V_n(x) = V(x)$ for $x \in K_n$ and $V_n, \nabla V_n$ are globally bounded;

(3) $(1/2)Ric - \overline{\nabla V_n} \ge (K-1)Id$ over M_n for all n.

On the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with the $\mathbb{R}^d$-valued Brownian motion W_t, let (U_t^n) be the solution of

$$dU_t^n = \sum_{i=1}^d H_i^n(U_t^n) \circ dW_t^i + \tilde{V}_n(U_t^n)dt$$

and $X_t^n = \Pi(U_t^n)$, where $H_i^n, i = 1, \cdots, d$ are the canonical horizontal vector fields on $O(M_n)$, $\tilde{V}_n$ is the horizontal lift of V_n.

Consider the stopping time

$$\tau_n := \inf\{t \ge 0; X_t \notin K_n\}$$

that satisfies $\mathbb{P}(\tau_n \uparrow +\infty) = 1$ by the non-explosion of our diffusion under condition (2.20). Then by the uniqueness of the SDE, we have with probability one,

$$U_t^n = U_t,\ X_t^n = X_t,\ \forall 0 \le t \le \tau_n.$$

Regarding f_i as function on $[0,T] \times M_n$ with the convention that $f(t,x) = 0$ for $x \notin K$ and $F(\gamma)$ as a functional on $W_{x_0}(M_n)$, we can apply Lemma 2.4.2 to get

$$F(X^n) - \mathbb{E}F(X^n) = \int_0^T \langle \mathbb{E}\left(\int_t^T (A_t^{(n)})^*[(A_s^{(n)})^*]^{-1}\nabla_s^{(n)} F(X^n)|\mathcal{B}_t\right), dW_t\rangle \tag{2.26}$$

where $\nabla_s^{(n)} F$ is the gradient w.r.t. L^2-norm on $W_{x_0}(M_n)$, and

$$M_t^n = \frac{1}{2} Ric_{U_t^n} - \overline{\nabla V_n}(U_t^n)$$
$$A_0^{(n)} = I,\ \frac{dA_t^{(n)}}{dt} = A_t^{(n)} M_t^n.$$

Step 2: convergence in (2.26).

At first, for the left-hand side , let us make a last localization on the functional F :

$$F_n(X) := \Phi\left(\int_0^{T\wedge\tau_n} f_1(s, X_s)ds\ ,\ \dots\ ,\ \int_0^{T\wedge\tau_n} f_m(s, X_s)ds\right).$$

It is clear that $\mathbb{P}(\tau_n \uparrow +\infty) = 1$ implies $F_n(X) \to F(X)$ a.s., remarking that

$$|F_n(X) - F(X)| \le \sum_{i=1}^m \|\partial_i \Phi\|_\infty \|f_i\|_\infty (T - T \wedge \tau_n) \to 0.$$

Since we have $F_n(X) = F_n(X^n)$, we obtain that $F(X^n) \to F(X)$ a.s. and consequently in $L^p(\mathbb{P})$ for all $p \in [1, +\infty)$. Thus the left-hand side of (2.26) converges to $F(X) - \mathbb{E}F(X)$ as desired.

For the right-hand side of (2.26), denoting $\mathbb{E}(\cdot|\mathcal{B}_t)$ by $\mathbb{E}_{\mathcal{B}_t}$, we have :

$$\begin{aligned}
& \int_0^T \langle \mathbb{E}_{\mathcal{B}_t}\left(\int_t^T (A_t^{(n)})^*[(A_s^{(n)})^*]^{-1}\nabla_s^{(n)} F(X^n)ds\right), dW_t\rangle \\
= & \int_0^T \langle \mathbb{E}_{\mathcal{B}_t}\left(\int_t^T (A_t)^*[(A_s)^*]^{-1}\nabla_s F(X)\ 1_{T<\tau_n}\right), dW_t\rangle \\
& + \int_0^T \langle \mathbb{E}_{\mathcal{B}_t}\left(\int_t^T (A_t^{(n)})^*[(A_s^{(n)})^*]^{-1}\nabla_s^{(n)} F(X^n)ds\ 1_{T\ge\tau_n}\right), dW_t\rangle \\
:= & (I)_n + (II)_n.
\end{aligned} \tag{2.27}$$

Letting $n \to +\infty$, we have obviously in $L^2(\mathbb{P})$,

$$(I)_n \to (I) := \int_0^T \langle \mathbb{E}_{\mathcal{B}_t} \left(\int_t^T (A_t)^* [(A_s)^*]^{-1} \nabla_s F(X) ds \right), dW_t \rangle$$

the desired r.h.s. Indeed, we have by the same proof as Lemma 2.3.2,

$$\|{A_t}^* [(A_s)^*]^{-1}\| \leq e^{-K(s-t)}, \ \forall s \geq t$$

and for the functionals in $\mathcal{D}$

$$|\nabla_s F(\gamma)| \leq \sum_{i=1}^m \|\partial_i \Phi\|_\infty \|\nabla f_i\|_\infty =: C,$$

so that

$$\begin{aligned} \mathbb{E}|(I)_n - (I)|^2 &= \mathbb{E} \int_0^T \left| \int_t^T (A_t)^* [(A_s)^*]^{-1} \nabla_s F(X) ds \ 1_{T \geq \tau_n} \right|^2 dt \\ &\leq \int_0^T \left| \int_t^T C e^{-(K)(s-t)} ds \right|^2 dt \ \mathbb{P}(\tau_n \leq T) \\ &\leq C^2 T^3 e^{2|K|T} \ \mathbb{P}(\tau_n \leq T) \to 0 \end{aligned}$$

We proceed in the same way for $(II)_n$, since by our assumption (3) in Step 1 and the proof of Lemma 2.3.2,

$$\|{A_t^{(n)}}^* [(A_s^{(n)})^*]^{-1}\| \leq e^{-(K-1)(s-t)}, \ \forall s \geq t$$

and also

$$|\nabla_s^{(n)} F(\gamma)| \leq \sum_{i=1}^m \|\partial_i \Phi\|_\infty \|\nabla f_i\|_\infty =: C$$

we have

$$\begin{aligned} \mathbb{E}(II)_n^2 &\leq \mathbb{E} \int_0^T \left| \int_t^T (A_t^{(n)})^* [(A_s^{(n)})^*]^{-1} \nabla_s^{(n)} F(X^n) ds \ 1_{T \geq \tau_n} \right|^2 dt \\ &\leq C^2 T^3 e^{2(|K|+1)T} \mathbb{P}(\tau_n \leq T) \to 0. \end{aligned}$$

Consequently letting n go to infinity in (2.26) for $F(X^n)$, we get the desired martingale representation for $F(X)$. $\square$

2.5 Some applications

Throughout this section let X be the diffusion generated by $\frac{1}{2}\Delta + V$ as given in §4 and the Bakry-Emery minoration condition (2.20) be satisfied with $K > 0$.

2.5.1 Hoeffding's type concentration implied by LSI

Let $g \in C_b^1([0,T] \times M)$ and consider the functional related to the central limit theorem:

$$F_T(\gamma) = \frac{1}{\sqrt{T}} \int_0^T (g(t,\gamma_t) - \mathbb{E}g(t,X_t))\,dt.$$

If $g \in C_0^\infty([0,T] \times M)$, then by the log-Sobolev inequality in Theorem 2.4.5 with $C(T) \le 1/K^2$ and by the Herbst method developed in Ledoux [17], we have

$$\mathbb{E}e^{\lambda F_T(X)} \le \exp\left(\frac{\lambda^2 \|\nabla_x g\|_\infty^2}{2K^2}\right), \ \forall \lambda \in \mathbb{R}.$$

Approximating $g \in C_b^1([0,1] \times M)$ by $g_n \in C_0^\infty([0,1] \times M)$, we see that the above inequality holds again for $g \in C_b^1([0,1] \times M)$. Thus by Chebychev inequality and an optimization over λ, we obtain in the standard way the following Hoeffding's type concentration inequality:

$$\mathbb{P}\left(F_T(X) > r\right) \vee \mathbb{P}\left(F_T(X) < -r\right) \le \exp\left(-\frac{r^2 K^2}{2\|\nabla_x g\|_\infty^2}\right), \ \forall r > 0. \tag{2.28}$$

2.5.2 Talagrand's T_2-transportation inequality and Tsirel'son's inequality

When treating the invariance principle (or functional central limit theorem) or functional moderate deviations ([9]), we have to study the functionals of type

$$F_{T,\delta}(X) := \frac{1}{\sqrt{T}} \sup_{0 \le s \le t \le s+\delta \le 1} \int_{Ts}^{Tt} (g(u,X_u) - \mathbb{E}g(u,X_u))du \tag{2.29}$$

where $g \in C_b^1([0,T] \times M)$. The concentration inequality of $F_{T,\delta}$ could be deduced from

Corollary 2.5.1. *Let $g \in C_b^1([0,T] \times M \to \mathbb{R}^m)$ and $Y_t = g(t,X_t)$. Then Y satisfies the log-Sobolev inequality on $L^2([0,T],\mathbb{R}^m)$:*

$$Ent(F^2(Y)) \le 2\frac{\|\nabla_x g\|_\infty^2}{K^2} \mathbb{E} \int_0^T |\nabla_t F|^2(Y) dt \tag{2.30}$$

for all functions $F \in C_b^1(L^2([0,T],\mathbb{R}^m))$, where ∇ is the gradient on $L^2([0,T],\mathbb{R}^m)$. In particular,

(a) the law $\mathbb{P}_Y$ of Y satisfies on $L^2([0,T],\mathbb{R}^m)$ the Talagrand T_2-transportation inequality with constant $\frac{\|\nabla_x g\|_\infty^2}{K^2}$, i.e.,

$$W_2^2(\mathbb{Q}, \mathbb{P}_Y) \le 2\frac{\|\nabla_x g\|_\infty^2}{K^2} Ent\left(\frac{d\mathbb{Q}}{d\mathbb{P}_Y}\right)$$

for all probability measures $\mathbb{Q}$ on $L^2([0,T],\mathbb{R}^m)$ such that $\mathbb{Q} \ll \mathbb{P}_Y$, where

$$W_2(\mathbb{Q},\mathbb{P}_Y) := \left(\inf_\pi \iint_{(L^2([0,T],\mathbb{R}^m))^2} \|h-\tilde{h}\|^2_{L^2([0,T],\mathbb{R}^m)} \pi(dh,d\tilde{h})\right)^{1/2},$$

is the Wasserstein L^2-distance between $\mathbb{Q}$ and $\mathbb{P}_Y$, here the infimum is taken over all couplings π of $(\mathbb{Q},\mathbb{P}_Y)$, i.e., over all the probability measures on $(L^2([0,T],\mathbb{R}^m))^2$ such that their marginal laws are respectively $\mathbb{Q}$ and $\mathbb{P}_Y$.

(b) (Tsirel'son's inequality) For any non-empty subset $A \subset L^2([0,T],\mathbb{R}^m)$,

$$G(Y) := \sup_{k\in A} \langle Y - \mathbb{E}Y, k\rangle_{L^2([0,T],\mathbb{R}^m)}$$

satisfies

$$\begin{aligned} &\mathbb{E}\exp\left(\frac{K^2}{\|\nabla_x g\|^2_\infty} \sup_{k\in A}\left[\langle Y-\mathbb{E}Y,k\rangle_{L^2([0,T],\mathbb{R}^m)} - \frac{1}{2}\|k\|^2_{L^2([0,T],\mathbb{R}^m)}\right]\right) \\ &\quad\le \exp\left(\frac{K^2}{\|\nabla_x g\|^2_\infty}\mathbb{E}G(Y)\right). \end{aligned}$$

(c) In particular for $F_{T,\delta}(X)$ given in (2.29), we have

$$\begin{aligned} &\mathbb{E}e^{\lambda F_{T,\delta}(X)} \le \exp\left(\lambda \mathbb{E}F_{T,\delta}(X) + \frac{\delta\lambda^2\|\nabla_x g\|^2_\infty}{2K^2}\right), \ \forall \lambda\in\mathbb{R}; \\ &\mathbb{P}\left(F_{T,\delta}(X) - \mathbb{E}F_{T,\delta}(X) > r\right) \le \exp\left(-\frac{r^2K^2}{2\delta\|\nabla_x g\|^2_\infty}\right), \ \forall r>0. \end{aligned}$$

Proof. At first for $F(h) = f(\langle h,e_1\rangle,\cdots,\langle h,e_n\rangle)$ where $e_i \in C^\infty([0,T],\mathbb{R}^m)$ and $f\in C^1_b(\mathbb{R}^n)$, then $F(Y)=\tilde{F}(X)$ where

$$\tilde{F}(\gamma) = f\left(\int_0^T e_1(t)g(t,\gamma_t)dt,\cdots,\int_0^T e_n(t)g(t,\gamma_t)dt\right)$$

belongs to our space $\mathcal{D}$ of test-functions if $g\in C_0^\infty([0,T]\times M)$. Noting that

$$\nabla_t\tilde{F}(X) = (\nabla_t F)(\cdots)U_t^{-1}\nabla_x g(t,X_t)$$

the log-Sobolev inequality (2.30) follows from Theorem 2.4.5. Approximating g by $g_n\in C_0^\infty([0,T]\times M)$, we obtain (2.30) for all such smooth cylindrical functions F.

Next for $F\in C_b^1(L^2([0,T],\mathbb{R}^m))$, let $(e_i)_{i\in\mathbb{N}} \subset C^\infty([0,T],\mathbb{R}^m)$ be an orthonormal basis of $L^2([0,T],\mathbb{R}^m)$ and $F_n(h) := F(P_n h)$ where P_n is the orthogonal projection to the subspace spanned by $(e_i)_{0\le i\le n}$. By the continuous differentiability of F,

$\nabla F_n(h) \to \nabla F(h)$ in $L^2([0,T],\mathbb{R}^m)$ for each h. Hence we obtain (2.30) for such F by the dominated convergence.

Both parts (a) and (b) are consequences of the log-Sobolev inequality (2.30), as shown by Bobkov-Gentil-Ledoux [7] (certainly they established them only on $\mathbb{R}^n$, but their arguments work for the actual Hilbert space setting).

Finally in the case of part (c), $m=1$, letting

$$A := \left\{ \frac{\lambda C}{\sqrt{T}} 1_{[Ts,Tt]};\ 0 \le s \le t \le s+\delta \le 1 \right\},\ C := \frac{\|\nabla_x g\|_\infty^2}{K^2}$$

we get part (c) from the Tsirel'son inequality in part (b). □

Bibliography

[1] S. Aida, Gradient estimates of harmonic functions and the asymptotics of spectral gaps on path spaces. *Interdiscip. Inform. Sci.* **2**, *no. 1, 75-84 (1996).*

[2] S. Aida, T. Masuda, I. Shigekawa, Logarithmic Sobolev inequalities and exponential integrability. *J. Funct. Anal.* **126**, *no. 1, 83-101 (1994).*

[3] S. Aida, D.Stroock, Moment estimates derived from Poincaré and logarithmic Sobolev inequalities. *Math. Res. Lett.* **1**, *no. 1, 75-86 (1994).*

[4] D. Bakry, *L'hypercontractivité et son utilisation en théorie des semigroupes,* Ecole d'Eté de Probabilités de Saint-Flour (1992), LNM 1581, Springer-Verlag 1994.

[5] D. Bakry, M. Emery. Diffusions hypercontractives. Seminaire de probabilités XIX, *Lecture Notes in Math, Volume* **1123**, *177-206 (1985).*

[6] S.G. Bobkov, F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163**, *1-28 (1999).*

[7] S. Bobkov, I. Gentil, M. Ledoux. Hypercontractivity of Hamilton-Jacobi equations. *J. Math. Pure Appl.,* **80**, *no. 7, 669-696 (2001).*

[8] M. Capitaine, E.P. Hsu and M. Ledoux. Martingale representation and a simple proof of logarithmic Sobolev inequality on path spaces. *Elect. Comm. Probab.,* **2**: *paper 7 (1997).*

[9] H. Djellout, A. Guillin, Moderate deviation for Markov chains with atom. *Stochastic Processes and Their Applications,* **95**, *203-217 (2001).*

[10] H. Djellout, A. Guillin, L. Wu, Transportation cost-information inequalities and application to random dynamical systems and diffusions, *Ann. Probab.,* **32**, *2702-2732 (2004).*

[11] H. Djellout, A. Guillin, L. Wu, Moderate deviations for non-linear functionals of moving average processes. Preprint 2004, to appear in *Ann. Inst. H. Poincaré.*

[12] S. Fang, Inégalité du type de Poincaré sur l'espace des chemins riemanniens, *CRAS,* **318**, *257-260 (1994).*

[13] Gentil, Inégalités de Sobolev logarithmiques et hypercontractivité en mécanique statistique et en EDP, *Thèse de Doctorat 2001, Université de Toulouse III.*

[14] E.P. Hsu, *Stochastic Analysis on Manifolds.* Graduate Studies in Mathematics, Volume 38, 2001.

[15] E.P. Hsu, Analysis on Path and Loop Spaces. *Probability theory and applications, Princeton, NJ, 1996, 277-347.*

[16] L. Gross, Logarithmic Sobolev inequalities, *Amer. J. Math.,* **97**, *1061-1083 (1975).*

[17] M. Ledoux. *Concentration of measure and logatithmic Sobolev inequalities.* Séminaire de Probabilités XXXIII. *Lecture Notes in Math. Springer*, 1709, 120-216 (1999).

[18] N. Ikeda, S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*, North Holland, 1989.

[19] D. Nualart. *The Malliavin Calculus and related Topics*, Springer Verlag, 1995.

[20] F. Otto, C. Villani Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality *J. Funct. Anal.,* **173**,*361-400 (2000).*

[21] C. Villani. *Topics in optimal transportation.* Grad. Stud. Math. (58), American Mathematical Society, 2003.

[22] F.Y. Wang. Transportation cost inequalities on path spaces over Riemannian manifolds. *Illinois J maths* **46**, *1197-1206 (2002).*

[23] L. Wu. A deviation inequality for non reversible Markov processes. *Ann. Inst. Henri Poincaré, Probabilités et Statistiques* **36**, *4, 435-445 (2000).*

[24] L. Wu and Z.L. Zhang, Talagrand's T_2-transportation inequality w.r.t. a uniform metric for diffusions, *Acta Math. Appl. Sinica, English Series, Vol.* **20**, *No.3, 357-364 (2004).*

Deuxième partie

Grandes Déviations au comportement ergodique pour des EDP stochastiques

Chapitre 3

Grandes déviations pour la mesure d'occupation d'une équation de Navier-Stokes ou Burgers stochastique

Dans ce chapitre, on présente une étude du comportement asymptotique des équations de Navier-Stokes 2D par des techniques de grandes déviations. On énonce aussi les principaux résultats des chapitres 4,5 et 6 concernant les équations de Navier-Stokes et Burgers.

3.1 Introduction

La compréhension de la turbulence hydrodynamique est depuis longtemps un axe de recherche majeur à la fois chez les ingénieurs, chez les physiciens et chez les mathématiciens. Avant les années 50 le salut venait souvent du recours à l'expérience et d'arguments phénoménologiques, comme la fameuse théorie de Kolmogorov [31]. La situation s'est transformée à partir des années 70 grâce aux progrès accomplis tant en matière de modélisation des écoulements, grâce aux équations de Navier-Stokes, qu'en capacité de traitement numérique, même si certains calculs semblent encore très coûteux.

Dans cette partie, on s'intéresse encore à ce que physiciens et ingénieurs qualifient *d'hypothèse d'ergodicité* (validée par l'expérience) qui explique la convergence en temps de moyennes de certaines quantités liées au fluide. Plus précisément, la vitesse de cette convergence probabiliste en un sens à définir, et les événements rares y afférents sont étudiés à l'aide de développements récents de la théorie des grandes déviations.

Il faut en premier lieu définir un cadre rigoureux dans lequel une telle convergence n'est plus une hypothèse mais peut être établie : l'ajout d'une force stochastique correcte *régularise* le système et permet d'obtenir l'existence et l'unicité d'un comportement asymptotique i.e d'une seule mesure invariante. On a alors le théorème ergodique qui donne ladite convergence.

Ce fait est central dans ce qui suit, mais ce n'est pas la seule justification à l'ajout d'une force aléatoire. Une telle perturbation peut aussi être vue comme un moyen de *tester* le système, pour comprendre son comportement hors de son état d'équilibre, comme Kolmogorov (encore lui) l'avait suggéré dans son approche de la turbulence. Enfin cette méthode est courante pour remplacer des bruits connus mais dont la prise en compte rendrait le modèle trop compliqué.

Un cadre mathématique général est d'abord introduit afin de souligner que l'intérêt du travail qui suit n'est pas restreint aux équations de la dynamique des fluides. En effet l'équation de Navier-Stokes 2D considérée ici constitue un exemple raisonnable d'équation stochastique non linéaire et dissipative incluant entre autres la plupart des équations dites de réaction-diffusion. Comme les résultats obtenus tiennent beaucoup au caractère dissipatif de l'équation et à la structure du bruit, cette approche devrait pouvoir s'appliquer à d'autres modèles aux calculs près.

Par exemple, l'équation de Burgers stochastique 1D est la forme asymptotique de nombreux systèmes non linéaires et dissipatifs. Elle modélise avec un certain succès certaines dynamiques des gazs, ou encore des phénomènes accoustiques. Son étude peut être traitée par des techniques analogues à celles développées dans ce chapitre avec un cadre un peu plus simple. Elles est insérée au chapitre 6 sous la forme d'un article.

A notre connaissance, l'étude des grandes déviations au comportement ergodique proposée ici est totalement nouvelle pour une Equation aux Dérivées Partielles Stochastique.

3.2 Le cadre

On présente brièvement le cadre des systèmes dynamiques Markoviens, et les grandes déviations associées.

3.2.1 Processus Stochastiques

Du fait de la perturbation aléatoire, et grâce à un certain niveau d'unicité, les solutions des équations considérées ne vivent pas chez les fonctions mais chez les processus de Markov. On commence donc par rappeler quelques définitions, en renvoyant au livre de Ethier et Kurz [19] pour une introduction complète aux processus de Markov.

Soit E un espace de Banach séparable de norme $|\cdot|$, et $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité complet. Une famille $\{X_t\}_{0\le t\le T}$ de variables aléatoires sur cet espace prenant leurs valeurs dans $(E,\mathcal{B}(E))$ est appelée processus stochastique. L'indice t est appelé le temps et E l'espace d'état du processus. Notre intérêt réside dans le comportement pour T grand des trajectoires du processus, c'est à dire des applications, $t\in[0,T]\to X(t,\omega)\in E$ pour chaque $\omega\in\Omega$.

Soit $\{\mathcal{F}_t\}$ une filtration de $\mathcal{F}$ (i.e une famille croissante de sous tribus). On dit que X est adapté à $\{\mathcal{F}_t\}$ si $\forall\ 0\le t\le T$, la variable aléatoire X_t est $\mathcal{F}_t$ mesurable. Dans le cas usuel où $\mathcal{F}_t=\sigma\{X_s\ :\ s\le t\}$, le processus adapté X_t est dit de Markov si

$$\mathbb{P}(X_t\in\Gamma\mid\mathcal{F}_s)=\mathbb{P}(X_t\in\Gamma\mid X_s),\quad\forall\ 0\le s\le t\le T,\quad\forall\Gamma\in\mathcal{B}(E).$$

Dans ce travail, le processus de Markov homogène $\{X_t\}_{0\le t\le T}$ décrit l'évolution aléatoire d'une particule dans l'espace d'état. On peut dans ce cas se demander quelle est la probabilité de trouver dans l'ensemble $\Gamma\in\mathcal{B}(E)$ à l'instant t, la particule partie du point $x\in E$ à l'instant initial. On dénote par $X(t,x)$ la position de cette particule pour souligner la condition initiale, et on introduit la fonction de transition,

$$P_t(x,\Gamma)\ :=\ \mathbb{P}(X(t,x)\in\Gamma)\ =\ \mathbb{P}_x(X_t\in\Gamma).$$

En fait, toute fonction de transition définit un semigroupe d'opérateurs linéaires sur l'ensemble $b\mathcal{B}(E)$ des fonctions mesurables et bornées sur E par

$$P_t\Psi(x)\ :=\ \int_E\Psi(y)P_t(x,dy)\ =\ \mathbb{E}\Psi\left(X(t,x)\right)\ =\ \mathbb{E}^x\Psi(X_t),\quad\forall\Psi\in b\mathcal{B}(E).\quad(3.1)$$

L'opérateur $P_t: b\mathcal{B}(E)\to b\mathcal{B}(E)$ ainsi défini est appelé *semigroupe de transition.* La propriété de semigroupe est donnée par l'identité de Chapman-Kolmogorov

$$P_{t+s}(x,\Gamma)=\int_E P_t(x,dy)P_s(y,\Gamma).$$

Dans le cas homogène, la propriété de Markov peut être formulée par

$$\mathbb{E}\left(\Psi(X_t)|\mathcal{F}_s\right)=P_{t-s}\Psi(X_s)\qquad\forall\Psi\in b\mathcal{B}(E).$$

Enfin, on rappelle qu'un semigroupe de Markov est dit

stochastiquement continu si $\lim_{t\to0}P_t\left(x,B(x,\delta)\right)=1,\forall x\in E,\delta>0$,

Feller si $P_t: C_b(E)\to C_b(E),\quad\forall t>0$,

fortement Feller si $P_t: b\mathcal{B}(E)\to C_b(E),\quad\forall t>0$, et

topologiquement irréductible si $P_t(x,\Gamma)>0,\forall t>0,\forall x\in E$ pour tout ouvert non vide $\Gamma\in\mathcal{B}(E)$.

3.2.2 Mesures invariantes et notions d'ergodicité

Les résultats suivants sont extraits du livre de Da Prato et Zabczyk [13]. Un semigroupe de transition P_t agit aussi sur l'ensemble $M_1(E)$ des mesures de probabilité sur E par l'intermédiaire de son semigroupe dual P_t^* :
$\forall t > 0$, $\forall \mu \in M_1(E)$, $\forall \Gamma \in \mathcal{B}(E)$,

$$P_t^* \mu(\Gamma) := \int_E P_t(x, \Gamma)\mu(dx) = \int_E \mathbb{P}_x(X_t \in \Gamma)\mu(dx) = \mathbb{P}_\mu(X_t \in \Gamma)$$

Plus précisément, $(P_t^*)_{t\geq 0}$ gouverne l'évolution de la distribution de X pour $t \geq 0$, i.e

$$\mathcal{L}(X_t) = P_t^* \mathcal{L}(X_0)$$

où $\mathcal{L}(Y) \in M_1(E)$ est la loi de la variable aléatoire Y à valeurs dans E.

Une mesure de probabilité $\mu \in M_1(E)$ est dite *invariante* ou *stationnaire* si c'est un point fixe du système dynamique défini par P_t^* sur $M_1(E)$, i.e $P_t^* \mu = \mu$. De manière équivalente, on a alors

$$\mathbb{E}^\mu f(X_t) \; = \; \int_E f(x)\mu(dx) \; := \; \mu(f), \qquad \forall t \geq 0.$$

D'après le critère de Krylov-Bogoliubov, pour assurer l'existence d'une mesure invariante il suffit de montrer que la famille de lois

$$\frac{1}{T}\int_0^T P_t^* \nu dt \tag{3.2}$$

est tendue pour une mesure initiale $\nu \in M_1(E)$ quelconque. En pratique, le cas $\nu = \delta_0$ est le plus fréquent.

Certaines mesures invariantes particulières donnent plus d'information sur le comportement asymptotique de la dynamique générée par P_t, il s'agit des mesures ergodiques dont une caractérisation pratique est donnée par le

Théorème 3.2.1. *Soit P_t un semi-groupe de Markov stochastiquement continu, et μ une mesure invariante pour P_t, $t \geq 0$. Les conditions suivantes sont équivalentes :*
(i) μ est ergodique
(ii) Pour tout $\phi \in L^2(E, \mu)$, on a la convergence suivante dans $L^2(E, \mu)$ et μ-p.s.

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^T P_s \phi ds = \int_E \phi(x)\mu(dx). \tag{3.3}$$

On déduit de ce résultat que deux mesures ergodiques distinctes d'un semigroupe de Markov stochastiquement continu sont singulières. Cependant, pour deux conditions initiales distinctes on peut observer une convergence vers deux mesures ergodiques distinctes. On parle alors de problèmes de transition de phase, qui rendent délicate l'étude de la vitesse de convergence.

On rappelle donc maintenant une caractérisation géométrique, qui souligne l'importance du problème de l'unicité de la mesure invariante.

Théorème 3.2.2. *Une mesure de probabilité est ergodique pour le semigroupe P_t si et seulement si il s'agit d'un point extremal de l'ensemble convexe de toutes les mesures de probabilité invariantes pour P_t.*
En particulier, si la mesure invariante est unique, elle est aussi l'unique mesure ergodique.

En bref, un cadre où la mesure invariante ergodique est unique est plus agréable pour étudier la vitesse de la convergence donnée par (3.3). Malheureusement, il n'y a pas de recette systématique pour montrer cette unicité. Récemment, dans le cas des équations de Navier-Stokes stochastique 2D, ce sont par exemple des méthodes de calcul de Malliavin ainsi qu'une analyse fine de la non-linéarité qui ont permis à Hairer, Mattingly et Pardoux [29, 37] de montrer cette unicité pour des bruits très dégénérés (i.e. une théorie d'hypoellipticité d'Hörmander infini-dimensionelle).

Cependant les outils récents de grandes déviations que nous utilisons requièrent certaines propriétés sur le semigroupe qui sont vérifiées si ce dernier est fortement Feller et topologiquement irréductible. Or, dans ce cas, le théorème suivant de Doob nous assure l'unicité.

Théorème 3.2.3. *Soit P_t, $t \geq 0$ un semi-groupe de Markov fortement Feller et topologiquement irréductible, μ une mesure de probabilité invariante pour P_t, alors :*
(i) μ est l'unique mesure de probabilité invariante pour P_t.
(ii) Pour tout $x \in E$, et tout $\Gamma \in \mathcal{B}(E)$,

$$\lim_{t\to\infty} P_t(x,\Gamma) = \mu(\Gamma).$$

(iii) μ est équivalente à toutes les mesures $P_t(x,\cdot)$, pour tout $x \in E$ et tout $t > 0$.

Sous ces conditions, la convergence est vraie pour toutes conditions initiales. On peut s'intéresser à la vitesse de convergence, et aux événements rares associés grâce à la théorie des grandes déviations, pour obtenir des résultats uniformes sur certains ensembles de conditions initiales.

3.2.3 Grandes déviations

Introduction et définitions

Pourquoi *grandes déviations*? Considérons le cas le plus simple, où $(X_n)_{n\in\mathbb{N}}$ est une suite de v.a.i.i.i.d dans $\mathbb{R}$, définies sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. Soit $m = \mathbb{E}(X_0)$ la moyenne commune des X_i et soit $\frac{S_n}{n} = \frac{1}{n}\sum_{k=1}^n X_k$ la moyenne empirique. On sait d'après la loi des grands nombres que,

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - m\right| > \varepsilon\right) = 0, \quad \forall\varepsilon > 0.$$

La théorie des grandes déviations s'intéresse à estimer, lorsqu'elle devient exponentiellement petite, la probabilité de l'évènement $\{|\frac{S_n}{n} - m| > \varepsilon\}$ donnant la déviation à la loi des grands nombres.

La naissance de cette théorie remonte aux travaux de Cramer en 1938. On peut diviser ses développements modernes en deux branches :

- *les petits paramètres*, comme dans l'ouvrage de Freidlin et Wentzell [27], lorsqu'on s'intéresse aux petites perturbations des systèmes dynamiques et
- *les grands paramètres*, liés à l'écart au comportement ergodique en temps grands. Deux ouvrages sont incontournables, celui de Deuschel-Stroock [17] et celui de Dembo-Zeitouni [15].

C'est dans la seconde catégorie que se situe notre propos : le but est de caractériser la vitesse exponentielle de la convergence ergodique en établissant un Principe de Grandes Déviations (PGD) pour la *mesure empirique* associée aux modèles considérés. De quoi s'agit il ?

Le premier résultat de ce type est le fameux théorème de Sanov (1957) qui traite le cas d'une suite de v.a.i.i.d $(X_n)_{n\in\mathbb{N}}$ à valeurs dans E, et donne un PGD pour la mesure

$$L_n := \frac{1}{n}\sum_{k=1}^{n}\delta_{X_k}$$

où δ_a est la mesure de Dirac en $a \in E$.

A chaque réalisation de la suite $(X_n)_{n\in\mathbb{N}}$ correspond une mesure de probabilité, de sorte que $L_n : \Omega \to M_1(E)$ est une variable aléatoire sur $M_1(E)$. En fait, L_n est appelée mesure empirique ou *mesure d'occupation* associée à la suite $(X_n)_{n\in\mathbb{N}}$. Cette dénomination est justifiée par le fait que $L_n(A)$ donne le temps moyen passé dans l'ensemble $A \in \mathcal{B}(E)$ par les n premiers termes de la suite.

Dans le cas du théorème de Sanov, on peut s'appuyer sur la loi forte des grands nombres pour obtenir la convergence de L_n vers la loi commune des X_k. En fait, on a $\mathbb{P}(L_n \in B) = \exp\left(-n \inf_{\nu\in B} H(\nu|\mu) + \circ(n)\right)$ pour $B \subset M_1(E)$ mesurable et $H(\nu|\mu)$ l'entropie relative (voir [15]). Mais pour avoir un comportement analogue dans le cas dépendant, les éléments d'ergodicité de la section précédentes sont cruciaux .

A partir de 1975, les travaux fondateurs de Donsker et Varadhan [18] dans le cas Markovien ont ouvert la voie à de très nombreux développements. Comme il est difficile d'établir une bibliographie sans oublier de contributions importantes, nous renvoyons à [17] contenant quelques extensions au temps continus, ou à [15] pour une liste de travaux sur le sujet antérieurs à 1998.

Quelle est la définition précise du PGD cherché ? Dans la suite, X_t est un processus de Markov qui admet des trajectoires continues à valeurs dans E (le rôle de X_t est joué par la solution d'une équation de Navier-Stokes stochastique et celui de E par un espace fonctionnel à préciser). On note encore $\mathbb{P}_x$ la loi sur $C(\mathbb{R}^+, E)$ du processus de Markov partant de $x \in E$. Si la condition initiale est une variable aléatoire de loi ν sur E, la loi du processus associé est $\mathbb{P}_\nu(\cdot) = \int_E \mathbb{P}_x(\cdot)\nu(dx)$.

La *mesure d'occupation* (ou *mesure empirique de niveau* 2) de ce processus est cette fois définie par

$$L_T := \frac{1}{T}\int_0^T \delta_{X_s} ds.$$

Il s'agit encore d'une variable aléatoire sur $M_1(E)$.

On rappelle que la topologie faible (notée w) sur $M_1(E)$ est la topologie de la convergence contre les fonctions $f \in C_b(E)$ continues et bornées sur E, alors que la *topologie forte* (notée τ) est celle de la convergence contre les fonctions boréliennes et bornées, $f \in b\mathcal{B}(E)$. C'est à dire que $\lambda_n \to \lambda$ dans $(M_1(E), \tau)$ si et seulement si on a $\lambda_n(f) \to \lambda(f), \forall f \in b\mathcal{B}(E)$.

L'approche proposée pour répondre à la question de la vitesse de convergence ergodique est donnée par la définition qui suit d'un bon PGD pour L_T, uniforme sur un ensemble de mesures initiales.

Définition 3.2.4. *Soit $\mathcal{M} \subset M_1(E)$ un ensemble de mesures initiales. On dit que la famille $\{\mathbb{P}_\nu(L_T \in \cdot)\}_{T\geq 0}$ suit un Principe de Grandes Déviations*

– *sur $M_1(E)$ equipé de la topologie forte τ,*
– *de fonction de taux J semi-continue inférieurement (s.c.i),*
– *uniformément sur l'ensemble de mesures initiales $\nu \in \mathcal{M}$*

si les trois propriétés suivantes sont vérifiées :

(a1) $J : M_1(E) \to [0, +\infty]$ est inf-compact pour la topologie τ, i.e $[J \leq a]$ est compact pour tout $a \geq 0$

(a2)(borne inférieure) pour tout τ-ouvert G dans $M_1(E)$,

$$\liminf_{T\to\infty} \frac{1}{T} \log \inf_{\nu\in\mathcal{M}} \mathbb{P}_\nu(L_T \in G) \geq -\inf_G J \tag{3.4}$$

(a3)(borne supérieure) pour tout τ-fermé F dans $M_1(E)$,

$$\limsup_{T\to\infty} \frac{1}{T} \log \sup_{\nu\in\mathcal{M}} \mathbb{P}_\nu(L_T \in F) \leq -\inf_F J. \tag{3.5}$$

Lorsqu'on cherche à établir un PGD, l'identification de la fonction de taux J peut être difficile. En pratique, elle est souvent obtenue à partir d'une fonction de taux classique (i.e donnée par un théorème type Schilder, Sanov ou Donsker-Varadhan) via un principe de contraction, résultat qui permet de déduire un PGD nouveau d'un PGD connu par action d'une fonction continue [15].

Le PGD de la définition précédente est bien plus précis que (3.3). Il permet par exemple de déduire une convergence exponentielle asymptotique quand $T \to \infty$ du type

$$\mathbb{P}_\nu\left(\left| \frac{1}{T}\int_0^T f(X_t)dt - \int_E f(x)\mu(dx) \right| \geq \delta \right) = \mathbb{P}_\nu\left(|L_T(f) - \mu(f)| \geq \delta\right) = e^{-Tc+\circ(T)}$$

pour $\nu \in \mathcal{M}$, $f \in b\mathcal{B}(E)$ et $\delta > 0$ où l'ensemble

$$F = \{\lambda \in M_1(E) \ \text{ t.q. } \ |\lambda(f) - \mu(f)| \geq \delta\}$$

est fermé pour la topologie τ. De plus ce PGD donne la vitesse exacte en cT où la constante c est connue comme l'infimum sur un certain ensemble de la fonction de taux J (qui dans notre cas ne s'annule qu'en l'unique mesure d'équilibre μ).

Le critère de récurrence hyper-exponentielle

Parmi les développements récents sur les Grandes Déviations, les travaux de Wu sur les processus de Markov essentiellement irréductibles en temps continu [44, 45, 46] contiennent différents critères pour établir des PGD comme celui de la définition 3.2.4. En particulier, une version du théorème ci-dessous est prouvée dans [46] (théorème 2.1) pour étudier le comportement d'un système Hamiltonien perturbé (une EDS).

Théorème 3.2.5. *Soit $\mathcal{M} \subset M_1(E)$ un ensemble de mesures initiales. On suppose que*

$$P_t \text{ est fortement Feller et topologiquement irréductible sur } E. \tag{3.6}$$

Si de plus $\forall \lambda > 0$, on peut trouver un compact $K \subset\subset E$, tel que les conditions d'intégrabilité

$$\sup_{\nu \in \mathcal{M}} \mathbb{E}^\nu e^{\lambda \tau_K} < \infty \qquad et \qquad \sup_{x \in K} \mathbb{E}^x e^{\lambda \tau_K^{(1)}} < \infty \tag{3.7}$$

soient vérifiées par les temps de retour définis par

$$\tau_K := \inf\{t \geq 0 \ t.q \ X_t \in K\} \quad et \quad \tau_K^{(1)} := \inf\{t \geq 1 \ t.q \ X_t \in K\},$$

alors la famille $\{\mathbb{P}_\nu(L_T \in \cdot)\}_{T \geq 0}$ vérifie le PGD de la définition 3.2.4
- *sur $M_1(E)$ équipé de la topologie forte τ,*
- *de bonne fonction de taux J, l'entropie de Donsker-Varadhan (3.10)*
- *uniformément sur l'ensemble de mesures initiales $\nu \in \mathcal{M}$.*

Quelques étapes de la preuve de ce résultat inspirées de [45, 46] sont insérées dans les chapitres 6 et 7 de ce document pour l'auto-contenance, l'objectif principal étant l'application aux modèles. Avant d'introduire l'équation de Navier-Stokes stochastique, on rappelle l'expression de l'entropie de Donsker-Varadhan J qui dirige le PGD ci-dessus. Les propriétés intéressantes concernant le minimum et l'ensemble de finitude de J sont étudiées dans les chapitres 4 et 6 consacrés aux équations.

La fonctionnelle d'entropie de Donsker-Varadhan.

Dans ce paragraphe, on suit Wu [46]. Afin d'introduire la fonctionnelle J associée, on dénote le processus de Markov X_t à trajectoires continues dans E par

$$\left(\Theta, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, (X_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E}\right)$$

et son semigroupe de noyau de transition par $(P_t(x,dy))_{t\geq 0}$, où $\Theta = C(\mathbb{R}^+, E)$ est l'espace des fonctions continues de $\mathbb{R}^+$ dans E muni de la topologie de la convergence sur les compacts ; pour tout $0 \leq s \leq t$, $\mathcal{F}_t^s = \sigma(X_u, s \leq u \leq t)$ est la filtration naturelle associée ; on convient que $\mathcal{F} = \sigma(X_u, 0 \leq u)$, $\mathcal{F}_t = \sigma(X_u, 0 \leq u \leq t)$ et on remarque enfin que $\mathbb{P}_x(X_0 = x) = 1$.

D'abord, la *mesure empirique de niveau* 3, c'est à dire au niveau du processus, est définie par

$$R_T := \frac{1}{T}\int_0^T \delta_{\theta_s X} ds \tag{3.8}$$

où $\theta_s : X \in \Theta \to \theta_s X \in \Theta$ avec $(\theta_s X)_t = X_{s+t}$ pour $t, s \geq 0$ sont les opérateurs de décalage (*shifts*) sur Θ. Ainsi R_t est un élément aléatoire de $M_1(\Theta)$, l'espace des mesures de probabilité sur Θ.

La fonctionnelle d'entropie de Donsker Varadhan de niveau 3 associée, notée $H : M_1(\Theta) \to [0, +\infty]$, est définie par

$$H(Q) := \begin{cases} \mathbb{E}^{\bar{Q}} h_{\mathcal{F}_1}\left(\bar{Q}_{X(-\infty,0]}; \mathbb{P}_{X(0)}\right) \text{ si } Q \in M_1^s(\Theta) \\ +\infty \text{ sinon} \end{cases} \tag{3.9}$$

où $M_1^s(\Theta)$ est le sous espace de $M_1(\Theta)$ composé des mesures qui sont de plus stationnaires. Dans l'expression (3.9), on note $\bar{Q}$ l'unique extension stationnaire de $Q \in M_1^s(\Theta)$ à l'espace $\bar{\Theta} := C(\mathbb{R}, E)$. La tribu $\mathcal{F}_n^m$ sur $\bar{\Theta}$ est alors définie $\forall m, n \in \mathbb{R}$ par $\mathcal{F}_n^m = \sigma(X(t); m \leq t \leq n)$. On note $\bar{Q}_{(-\infty,t]}$ une version régulière de la distribution conditionnelle de $\bar{Q}$ connaissant $\mathcal{F}_t^{-\infty}$. Puis, *l'entropie relative* $h_{\mathcal{G}}(\nu, \mu)$ (ou *information de Kullback*) de ν par rapport à μ, restreinte à la tribu $\mathcal{G}$ est donnée par

$$h_{\mathcal{G}}(\nu, \mu) := \begin{cases} \int \frac{d\nu}{d\mu}|_{\mathcal{G}} \log\left(\frac{d\nu}{d\mu}|_{\mathcal{G}}\right) d\mu, \text{ si } \nu << \mu \text{ sur } \mathcal{G} \\ +\infty \text{ sinon }. \end{cases}$$

On reconnait une forme analogue à l'entropie introduite dans la partie consacrée aux Inégalités de Sobolev Logarithmiques : l'estimation de cette quantité peut être réalisée à l'aide d'inégalités fonctionnelles.

Enfin *l'entropie de Donsker-Varadhan* notée $J : M_1(E) \to [0, \infty]$ (ou fonctionnelle d'entropie de niveau 2) qui intervient dans le critère précédent s'obtient par projection de la fonctionnelle de niveau 3 :

$$J(\beta) = \inf\{H(Q) \mid Q \in M_1^s(\Theta) \text{ et } Q_0 = \beta\}, \quad \forall \beta \in M_1(E), \tag{3.10}$$

où $Q_0(\cdot) = Q(X(0) \in \cdot)$ est la loi marginale en temps $t = 0$ (cette projection est en fait liée au principe de contraction).

Le Théorème 3.2.5 permet de décrire la convergence exponentielle, mais l'expression de la vitesse exacte donnée par J est plus qualitative que quantitative. En effet, $J(\nu)$ n'admet une forme explicite que dans le cas où l'unique mesure invariante μ est connue et le processus de Markov X est symétrique par rapport à μ. Dans les autres cas, on peut parfois obtenir des estimations sur $J(\nu)$, par exemple à l'aide d'inégalités fonctionnelles comme l'inégalité de Sobolev logarithmique, ou bien trouver des exemples de mesures ν pour lesquelles $J(\nu)$ peut être calculée.

3.3 Les équations de Navier-Stokes

Les équations de Navier-Stokes que nous étudions décrivent le mouvement d'un fluide visqueux incompressible et homogène. Le cadre fonctionnel et le traitement de ces équations dépendent de la dimension de l'espace. Nous nous restreignons au cas de la dimension 2 car une unique solution est alors définie pour des temps arbitrairement grands et on peut s'intéresser au problème du comportement asymptotique : déterminer quel régime permanent sera observé après une courte période initiale.

3.3.1 Modélisation

Soit D un domaine borné de $\mathbb{R}^2$ de frontière régulière ∂D. L'évolution de la vitesse $u(t,\xi)$ du fluide se trouvant en $\xi \in D$ au temps t est donnée par les équations de Navier-Stokes : l'équation de conservation du moment cinétique s'écrit

$$\frac{du(t,\xi)}{dt} + (u(t,\xi).\nabla)u(t,\xi) - \nu\Delta u + \nabla P(t,\xi) = g(t,\xi) \quad , \quad t \geq 0, \quad \xi \in D, \quad (3.11)$$

et la condition d'incompressibilité

$$\text{div } u(t,\xi) = 0 \quad , \quad t \geq 0 \quad , \quad \xi \in D \; .$$

On considère de plus des conditions aux bords de type non-glissement

$$u(t,\xi) = 0 \quad , \quad t \geq 0 \quad , \quad \xi \in \partial D \qquad (3.12)$$

et le champ des vitesses initiales est donné par

$$u(0,\xi) = u_0(\xi) \quad , \quad \xi \in D.$$

Le coefficient $\nu > 0$ est la viscosité du fluide (plus petite pour l'eau que pour le mazout...). Ici le fluide est supposé homogène (à densité constante), sinon il faut rajouter une équation de conservation de la masse. On peut aussi remplacer les conditions de Dirichlet homogènes (3.12) par d'autres conditions aux bords (voir la section 3.7.1 pour le cas périodique).

Pour une introduction complète à la théorie Mathématique des équations de Navier-Stokes, on renvoie aux livres de Temam [42, 43]. Le cadre présenté ci dessous ainsi que des formulations faibles pour différents types de conditions limites sont détaillés dans ces ouvrages.

3.3.2 Vers une formulation faible

En suivant [42], on introduit quelques espaces fonctionnels nécessaires. Soit $\mathcal{C}_0^\infty(D)$ l'espace des fonctions à valeurs réelles qui sont $\mathcal{C}^\infty$ à support compact dans D, et soit

$$\mathcal{V} = \{u = (u_1, u_2) \in (\mathcal{C}_0^\infty(D))^2 \quad / \quad \text{div } u = 0\}.$$

On dénote par H (resp. par V) sa fermeture dans l'espace produit $\left[L^2(D)\right]^2$ (resp. dans le produit $\left[H^1(D)\right]^2$ de l'espace des fonctions L^2 de dérivée L^2). La régularité du domaine permet de montrer l'identification

$$\begin{array}{rcl} H & = & \{u \in [L^2(D)]^2 \text{ t.q. div } u = 0, \quad \gamma_\eta(u) = 0 \text{ sur } \partial D\} \\ V & = & \{u \in [H_0^1(D)]^2 \text{ t.q. div } u = 0\} \end{array}$$

où $\gamma_\eta(u) = 0$ est un opérateur de trace qui coincide avec $u \cdot \eta$ pour u régulière, η étant la normale extérieure à ∂D (voir le chapitre 1 de [42]).

L'espace H est équipé du produit scalaire $\langle\ ,\ \rangle$ et de la norme $|\cdot|$ hérités de $\left[L^2(D)\right]^2$. De façon usuelle, on identifie H avec son dual H' de sorte que $V \subset H \subset V'$, et on note encore $\langle\ ,\ \rangle$ la dualité entre V' et V.

La procédure classique consiste à appliquer à l'équation (3.11) la projection orthogonale P_{div} de $\left(L^2(D)\right)^2$ sur H. Ceci a pour effet de supprimer la pression et les composantes irrotationnelles apparaissant dans la décomposition de Helmholtz de u (voir [42]). On obtient ainsi une nouvelle équation dans V' qui s'écrit

$$\frac{d}{dt}u(t) + Au(t) + B\left(u(t), u(t)\right) = f(t) \quad , \quad t \geq 0, \quad \xi \in D, \tag{3.13}$$

où A est l'opérateur de Stokes défini par

$$Au = -\nu P_{div} \Delta u \quad , \quad \forall u \in D(A) = \left[H^2(D)\right]^2 \cap V, \tag{3.14}$$

la forme bilinéaire $B : V \times V \to V'$ est définie par

$$\langle B(u,v), z\rangle := \int_D z(\xi) \cdot (u(\xi) \cdot \nabla) v(\xi) d\xi, \quad \forall u, v, z \in V \tag{3.15}$$

et $f(t)$ est la composante à divergence nulle de la force extérieure $g(t)$. Cette formulation faible due à Leray dans ses travaux de 1934 joue un rôle fondamental dans la théorie mathématique des équations de NS.

3.3.3 Les opérateurs A et B

L'opérateur A défini par (3.14) est linéaire non borné sur H et son inverse A^{-1}, qui donne la solution du problème de Stokes, est un opérateur auto-adjoint compact sur H. Des résultats classiques de théorie spectrale permettent d'en déduire l'existence d'une suite croissante

$$0 < \lambda_1 \leq \lambda_2 \leq ..., \qquad \lambda_k \to \infty$$

de valeurs propres de A ainsi qu'une base orthonormale de H formée de vecteurs propres $e_1, e_2, ..$ tels que $Ae_k = \lambda_k e_k$ avec $e_k \in D(A)$. On choisit systèmatiquement dans la suite cette base des (e_k) pour H. Dans le cas où D est un rectangle, ou si l'on travaille avec des conditions périodiques, les e_k sont des fonctions trigonométriques.

A partir de la décomposition spectrale précédente, on peut définir les puissances fractionnaires A^α et leurs domaines respectifs $D(A^\alpha)$ qui sont des sous espaces fermés de $[H^{2\alpha}(D)]^2$. De plus, la norme $|u|_\alpha := |A^\alpha u|$ sur $D(A^\alpha)$ est équivalente à la norme induite par l'espace de Sobolev $[H^{2\alpha}(D)]^2$. En particulier, l'espace V qui coïncide avec $D\left(A^{\frac{1}{2}}\right)$ est muni de la norme $|u|_V = |A^{\frac{1}{2}}u|$. Pour tout $\alpha > 0$, on a aussi

$$|u| \leq \frac{1}{{\lambda_1}^\alpha}|A^\alpha u|. \tag{3.16}$$

D'après Métivier [38], le comportement asymptotique des valeurs propres de A en dimension d est donné par $\lambda_k \sim \lambda_1 k^{\frac{2}{d}}$, ce qui signifie dans le cas bidimensionnel que l'on peut trouver $c, C \geq 0$ tels que l'on ait

$$c \leq \frac{\lambda_k}{k} \leq C, \tag{3.17}$$

pour k suffisament grand. Ainsi, l'injection $D(A^\alpha)$ dans $D\left(A^\beta\right)$ est dense et compacte pour $\alpha > \beta \geq 0$.

Enfin, l'opérateur $-A$ engendre un semigroupe analytique borné e^{-tA} et pour tout $\alpha \geq 0$, il existe une constante positive M_α vérifiant

$$\|A^\alpha e^{-tA}\|_{L(H)} \leq \frac{M_\alpha}{t^\alpha} \tag{3.18}$$

où $\|\cdot\|_{L(H)}$ est la norme usuelle des opérateurs linéaires sur H.

La forme bilinéaire B prend ses valeurs dans V'. En effet, pour tout $u, v, z \in V$, l'injection continue de H_0^1 dans L^4 implique que

$$\begin{aligned} |\langle B(u,v), z\rangle| &= \left|\int_D z(\xi)\cdot(u(\xi)\cdot\nabla)v(\xi)d\xi\right| \\ &\leq \|u\|_{[L^4(D)]^2}\ \|\nabla v\|_{[L^2(D)]^2}\ \|z\|_{[L^4(D)]^2} \end{aligned}$$

$$\leq \quad C|u|_V \ |v|_V \ |z|_V.$$

On sait aussi grâce à la condition d'incompressibilité que pour u, v, z dans V,

$$\langle B(u,v), v\rangle = 0 \quad , \quad \langle B(u,v), z\rangle = -\langle B(u,z), v\rangle. \tag{3.19}$$

En particulier, comme on a l'inclusion continue $D\left(A^{\frac{1}{4}}\right) \subset [L^4(D)]^2$, il vient

$$|\langle B(u,u), z\rangle| = |-\langle B(u,z), u\rangle| \leq C|A^{\frac{1}{2}}z| \ |A^{\frac{1}{4}}u|^2. \tag{3.20}$$

3.3.4 A propos du comportement des solutions de (3.13)

En dimension 2, étant données une force $f \in L^2(0,T;V')$ et une condition initiale $u_0 \in H$, l'équation (3.13) admet une unique solution vérifiant

$$u \quad \in \quad C([0,T];H) \cap L^2(0,T;V).$$

On observe parfois, après un certain laps de temps, la convergence du système vers un état d'équilibre, qui doit donc être une solution stationnaire du système. Or, on sait montrer l'existence de solution stationnaire, mais on n'en connaît pas le nombre.

Le seul résultat connu lorsque $T \to \infty$ est prouvé pour une viscosité ν suffisamment grande par rapport à l'amplitude $|f|$ de la force. L'ensemble des solutions stationnaires est alors réduit à un point, et le système converge vers cette configuration d'équilibre.

Mais pour les grandes vitesses ou les petites viscosités il y a beaucoup de solutions stationnaires, et le comportement asymptotique du système est plus compliqué. C'est un problème général avec les systèmes dynamiques non linéaires, la non linéarité impliquant un comportement chaotique, même dans des cas plus simples que Navier-Stokes. Pourtant la dissipativité implique l'existence d'ensembles absorbants et d'un attracteur compliqué mais dont on sait qu'il est fini-dimensionnel (voir [43]).

Cette complexité et le théorème ergodique sont la base de la dynamique des fluides statistique, qu'on peut comparer à la mécanique statistique qui traite les grands nombres de particules.

3.4 La perturbation aléatoire

Nous présentons maintenant le type de perturbation aléatoire ajoutée à l'équation (3.13). La référence [12] contient une présentation détaillée des objets introduits.

3.4.1 Wiener cylindrique et Q-Wiener

On appelle *processus de Wiener cylindrique* dans H le processus donné par la série

$$W(t) := \sum_{j\geq 1} \beta_j(t) e_j \tag{3.21}$$

où les β_j sont des mouvements Browniens réels indépendants sur $(\Omega, \mathcal{F}, \mathbb{P})$. Cette série ne converge pas dans H, et ne possède pas les propriétés (i) à (iv) ci dessous d'un vrai processus de Wiener sur H.

On appelle *opérateur à noyau*, ou *opérateur à trace sur* H et on note $Q \in L_1(H)$ un opérateur linéaire $Q : H \to H$ vérifiant $\sum_{j\geq 1} \langle Qe_j, e_j\rangle < \infty$. Si $Q \in L_1(H)$ est un opérateur symétrique positif à noyau, on appelle *Q-processus de Wiener* ou *Q-Wiener sur* H un processus stochastique $W^Q(t)$ vérifiant

(i) $W^Q(0) = 0$ $\mathbb{P}$-p.s

(ii) W^Q a des trajectoires continues $\mathbb{P}$-p.s

(iii) W^Q est à accroissements indépendants

(iv) la loi de $W^Q(t) - W^Q(s)$ est une gaussienne centrée de covariance $(t-s)Q$, i.e $\mathcal{L}(W^Q(t) - W^Q(s)) = \mathcal{N}(0, (t-s)Q)$.

Le processus $W^Q(t)$ est donc un processus Gaussien dans H vérifiant $\mathbb{E}W^Q(t) = 0$ et $cov(W^Q(t)) = tQ$.

Dans le cas du processus de Wiener cylindrique, $Q = I_d \notin L_1(H)$. On montre cependant que la série (3.21) définit bien un Q_1-Wiener, pas dans H mais plutôt dans un espace H_1 plus grand tel que l'inclusion de H dans H_1 soit Hilbert-Schmidt, i.e $\sum_{k\geq 1} |e_k|^2_{H_1} < \infty$. En dimension 2 par exemple, d'après (3.17), on sait que l'inclusion $H \subset D(A^{-(\frac{1}{2}+\varepsilon)})$ est Hilbert-Schmidt, et donc la série (3.21) définit un Q_1-Wiener dans tout espace H_1 vérifiant $D(A^{-(\frac{1}{2}+\varepsilon)}) \subset H_1$.

Mais l'exemple qui nous intéresse plus particulièrement est le suivant. On rappelle que l'opérateur linéaire $G : H \to H$ est *Hilbert-Schmidt*, noté $G \in L_2(H)$, si $\sum_{k\geq 1} |Ge_k|^2 < \infty$. La quantité précédente est une norme sur cet espace notée $\|G\|_{L_2(H)}$. Il est clair dans ce cas que l'opérateur symétrique positif GG^* est à noyau. On a en particulier le

Lemme 3.4.1. *Si l'opérateur linéaire borné $G : H \to H$ vérifie pour un $\varepsilon > 0$,*

$$Im(G) \subset D\left(A^{\frac{1}{2}+\varepsilon}\right) \tag{3.22}$$

où $Im(G)$ est l'image de G, alors l'opérateur symétrique positif $Q := GG^*$ est à trace, i.e $Q \in L_1(H)$.

De plus, si $W(t)$ est le processus de Wiener cylindrique (3.21), alors $GW(t)$ est un Q-Wiener sur H de covariance $Q := GG^*$.

Démonstration. D'après (3.17), l'opérateur $A^{-(\frac{1}{2}+\varepsilon)}$ est Hilbert-Schmidt. De plus, l'opérateur $A^{(\frac{1}{2}+\varepsilon)}G$ est linéaire et borné sous la condition (3.22). Donc l'opérateur $G = A^{-(\frac{1}{2}+\varepsilon)}\left(A^{(\frac{1}{2}+\varepsilon)}G\right)$ est aussi Hilbert-Schmidt comme composition d'un opérateur $L_2(H)$ avec un opérateur de $L(H)$. □

3.4.2 La convolution stochastique

Etant donné un Q-Wiener sur H, on peut construire l'intégrale stochastique pour les intégrandes qui sont des processus dans l'espace des opérateurs linéaires $L(H)$, alors que s'il s'agit d'un processus de Wiener cylindrique on doit se restreindre aux intégrandes qui sont des processus à valeurs dans l'ensemble $L_2(H)$ des opérateurs de Hilbert-Schmidt (voir [12]).

Dans cette section on considère l'intégrale stochastique

$$W_A(t) := \int_0^t e^{-(t-s)A} G dW(s) \tag{3.23}$$

où A est l'opérateur de Stokes défini par (3.14) et $W(t)$ est un processus de Wiener cylindrique. Le processus stochastique ainsi défini est appelé *convolution stochastique*. C'est en fait un processus type Ornstein-Uhlenbeck en dimension infinie solution de l'équation

$$dW_A(t) + AW_A(t)dt = GdW(t) \quad , \quad W_A(0) = 0. \tag{3.24}$$

Il est étudié en détails dans [12], le résultat suivant présente quelques propriétés utiles pour l'étude de l'équation NS-2D.

Proposition 3.4.2. *(i) Si $\int_0^T \|e^{-rA}G\|^2_{L_2(H)} dr < \infty$, alors le processus $W_A(t)$ est un processus Gaussien, continu en moyenne quadratique admettant une version prévisible. Sa loi est une mesure Gaussienne centrée sur $L^2(0,T;H)$.*

(ii) De plus, si $Im(G) \subset D(A^{\beta+\varepsilon})$ pour un $\beta > 0$, alors c'est un processus Gaussien à trajectoires continues dans $D(A^\beta)$.

(iii) Enfin si l'inclusion $Im(G) \subset D(A^{\beta+\varepsilon})$ est dense, alors la loi de W_A remplit tout l'espace $C_0([0,T], D(A^\beta))$ des fonctions continues et nulles en $t = 0$, i.e $\overline{supp}\mathcal{L}(W_A) = C_0([0,T], D(A^\beta))$.

Démonstration. On renvoie à [12] pour une preuve détaillée de (i) et (ii). L'idée est la suivante : si $Im(G) \subset D(A^{\beta+\varepsilon})$, on a alors

$$\begin{aligned}
\mathbb{E}|A^\beta W_A(t)|^2 &= \mathbb{E}\left|\int_0^t A^\beta e^{-(t-s)A} G dW(s)\right|^2 \\
&= \mathbb{E}\left|\int_0^t e^{-(t-s)A} A^{\frac{1-\varepsilon}{2}} (A^{-\frac{1+\varepsilon}{2}} A^{\beta+\varepsilon} G) dW(s)\right|^2
\end{aligned}$$

$$\begin{aligned} &= \int_0^t \|e^{-(t-s)A} A^{\frac{1-\varepsilon}{2}} (A^{-\frac{1+\varepsilon}{2}} A^{\beta+\varepsilon} G)\|^2_{L_2(H)} ds \\ &\leq C \int_0^t \|e^{-(t-s)A} A^{\frac{1-\varepsilon}{2}}\|^2_{L(H)} ds \end{aligned}$$

puisque $\|A^{-\frac{1+\varepsilon}{2}} A^{\beta+\varepsilon} G\|_{L_2(H)} \leq C$ vu la preuve du lemme précédent, et que

$$\|MN\|_{L_2(H)} \leq \|M\|_{L(H)} \|N\|_{L_2(H)}, \quad \forall M \in L(H), \forall N \in L_2(H).$$

Ainsi, avec (3.18), on obtient

$$\mathbb{E}|A^\beta W_A(t)|^2 \leq CM \int_0^t \frac{ds}{(t-s)^{1-\varepsilon}} < \infty.$$

Puis, pour s'assurer que le processus $W_A(\cdot)$ est à trajectoires continues dans $D(A^\beta)$, il suffit de montrer d'après le théorème 5.9 de [12] que pour un $\gamma > 0$, on ait

$$\int_0^T r^{-\gamma} \|A^\beta e^{-rA} G)\|^2_{L_2(H)} dr.$$

D'après le calcul précédent ceci est vrai pour $0 < \gamma < \varepsilon$. Enfin, la partie (iii) est un résultat de Maslowski [34]. □

Soit $a \geq 0$; il est parfois utile de considérer l'équation suivante de Ornstein-Uhlenbeck modifiée

$$dW_A^a(t) + AW_A^a(t)dt + aW_A^a(t)dt = GdW(t). \tag{3.25}$$

Si $Im(G) \subset D(A^{\beta+\varepsilon})$, cette équation a une solution stationnaire donnée par

$$W_A^a(t) := \int_{-\infty}^t e^{-(t-s)(A+a)} GdW(s) \tag{3.26}$$

qui est un processus Gaussien continu dans $D(A^\beta)$. L'intérêt technique est que l'on peut contrôler les moments d'ordre pair de ce processus. En effet, pour $p \geq 1$, on obtient par un calcul analogue au précédent

$$\begin{aligned} \mathbb{E}|A^\beta W_A^a(t)|^{2p} &\leq C \left(\mathbb{E}|A^\beta W_A^a(t)|^2\right)^p \\ &\leq C \left(\int_0^\infty \frac{e^{-2ra}}{r^{1-\varepsilon}} dr\right)^p \to 0 \text{ quand } \quad a \to \infty. \end{aligned}$$

Cette propriété est par exemple utilisée pour montrer l'existence d'une mesure invariante.

3.5 Le problème stochastique associé

Nous adoptons la formulation du problème stochastique introduite par Flandoli [24] qui permet de traiter des bruits relativement généraux. Nous présentons dans cette section certains résultats connus concernant l'existence et la régularité des solutions, ainsi que du semigroupe de transition associé. Ces résultats proviennent de travaux de Ferrario, Flandoli, Goldys et Maslowski (voir [22, 24, 25, 28]). Les principales idées des preuves sont données.

3.5.1 Formulation du problème stochastique

Le problème stochastique correspondant à l'équation (3.13) s'écrit

$$dX(t) + AX(t)dt + B\left(X(t), X(t)\right)dt = fdt + GdW(t)\ ;\ X(0) = x \tag{3.27}$$

où $G \in L(H)$ et $W(t)$ est un processus de Wiener cylindrique sur un espace de probabilités $(\Omega, \mathcal{F}, \mathbb{P})$ fixé. Comme on s'intéresse au comportement en temps grand, la force extérieure est supposée stationnaire afin d'avoir un système autonome : f ne dépend pas de t et le bruit $GdW(t)$ est blanc en temps.

Définition 3.5.1. *Un processus stochastique progressivement mesurable*

$$X \in C([0,T], H) \cap L^2\left(0, T; D(A^{\frac{1}{4}})\right) \quad \mathbb{P} - p.s$$

est appelé solution généralisée de l'équation (3.27) sur l'intervalle $[0,T]$ *si* $X(t)$ *vérifie*

$$\begin{aligned} & \langle X(t), y\rangle + \int_0^t \langle X(s), Ay\rangle ds - \int_0^t \langle B\left(X(s), y\right), X(s)\rangle ds \\ = \ & \langle x, y\rangle + t\langle f, y\rangle + \langle GW(t), y\rangle \end{aligned}$$

$\mathbb{P}$*-p.s, pour tout* $t \in [0,T]$ *et* $y \in D(A)$.

Cette définition introduite par Flandoli [24] a un sens : d'après (3.20) la régularité $X(t) \in L^2\left([0,T], D(A^{\frac{1}{4}})\right)$ est suffisante pour que le terme intégral en B soit bien défini et de plus cette représentation correspond bien à l'équation (3.27).

3.5.2 Existence et régularité des solutions

On suppose à partir de maintenant que (3.22) est vérifiée, i.e

$$Im(G) \subset D(A^{\frac{1}{2}+\varepsilon}). \tag{3.28}$$

En fait, Flandoli [24] a montré l'existence sous des conditions plus faibles. Cependant ceci est l'hypothèse du lemme 3.4.1, qui est central dans notre étude. On présente dans cette section et la suivante les versions de résutats connus qui correspondent à cette condition.

Théorème 3.5.2. *(i) Pour tout $T > 0$, toute condition initiale $x \in H$, et toute force $f \in D(A^{-\frac{1}{2}})$ il existe une unique solution généralisée à l'équation (3.27) vérifiant $\mathbb{P}$-p.s*

$$X \in C([0,T], H) \cap L^2\left(0, T; D(A^{\frac{1}{2}})\right)$$

et

$$X - W_A \in L^2\left(0, T; D(A^{\frac{1}{2}})\right).$$

De plus, X est un processus de Markov dans H.

(ii) Soit $0 < \alpha < \frac{1}{2}$. Si la condition initiale vérifie $x \in D(A^{\alpha})$ et la force $f \in D(A^{-\frac{1}{2}+\alpha})$ alors la solution X vérifie $\mathbb{P}$-p.s

$$X \in C([0,T], D(A^{\alpha})) \cap L^2\left(0, T; D(A^{\frac{1}{2}})\right) \cap L^{\frac{4}{1-2\alpha}}\left(0, T; D(A^{\frac{1}{4}+\frac{\alpha}{2}})\right)$$

et

$$X - W_A \in L^2\left(0, T; D(A^{\frac{1}{2}+\alpha})\right).$$

(iii) Enfin, si la force et la condition initiale vérifient $x, f \in H$, alors pour tout $t_0 > 0$ et tout $0 < \alpha < \frac{1}{2}$, la solution vérifie $\mathbb{P}$-p.s

$$X \in C([t_0, T], D(A^{\alpha})).$$

La preuve du résultat précédent est classique. L'idée de départ est la suivante : en soustrayant l'équation (3.24) à l'équation (3.27), on vérifie que le processus $V(t) := X(t) - W_A(t)$ satisfait à

$$dV(t) + AV(t)dt + B\left(V(t) + W_A(t), V(t) + W_A(t)\right)dt = f dt \; ; \; V(0) = x \tag{3.29}$$

Cette équation est étudiée en fixant une trajectoire $W_A(\omega)$, pour ω dans un ensemble de mesure plein où W_A possède la bonne régularité. On utilise alors les méthodes déterministes usuelles : majorations a priori, et projections de Galerkin.

Cette technique a été introduite dans le travail fondateur de Bensoussan et Temam [4], mais eux considéraient $V_{BT}(t) := X(t) - GW(t)$ ce qui leur imposait une plus forte hypothèse sur le bruit, puisque les trajectoires de GW sont beaucoup moins régulières que celles de W_A. En effet, l'hypothèse minimale pour prouver l'existence d'une solution généralisée dans [24], est $Im(G) \subset D(A^{\frac{1}{4}+\varepsilon})$, afin que W_A prennent ses valeurs dans $D(A^{\frac{1}{4}})$ alors que pour pouvoir définir une solution comme dans [4], il faut que le processus $GW(t)$ lui-même prenne ses valeurs dans $D(A^{\frac{1}{4}})$, une condition suffisante étant $Im(G) \subset D(A^{\frac{3}{4}+\varepsilon})$ (voir [36]).

A partir de l'équation vérifiée par $V(t,x_1)-V(t,x_2)$, on obtient avec l'estimation d'énergie classique et le lemme de Gronwall

$$Sup_{0\le t\le T}|V(t,x_1)-V(t,x_2)| \le C|x_1-x_2| \exp\left(\int_0^T |V(s,x_1)+W_A(s)|_{L^4}^4 ds\right).$$

La regularité requise à droite de cette estimation est l'obstacle jusqu'à présent infranchissable pour l'unicité en dimension 3. Par contre, comme $D(A^{\frac{1}{4}}) \subset L^4$, et vu la régularité pour W_A, on obtient le contrôle

$$Sup_{0\le t\le T}|V(t,x_1)-V(t,x_2)| \le C|x_1-x_2| \quad \mathbb{P}-p.s \tag{3.30}$$

pour une variable aléatoire C p.s finie, qui implique l'unicité du résultat précédent.

On peut trouver les détails des ces preuves dans la littérature. Les parties (i) et (ii) sont des cas particuliers du résultat de Flandoli [24]. Pour le (ii) voir aussi le travail de Ferrario [22] en ce qui concerne la régularité dans $L^{\frac{4}{1-2\alpha}}\left(0,T;D(A^{\frac{1}{4}+\frac{\alpha}{2}})\right)$ dont on se sert dans la preuve de l'irréductibilité. Enfin le (iii) est bien détaillé dans le travail de Flandoli et Maslowski [25].

3.5.3 Semigroupe associé et mesure invariante

On considère maintenant le semigroupe P_t associé au processus de Markov X solution de l'équation (3.27). Grâce à l'estimation (3.30) et au théorème de convergence dominée, on peut montrer que P_t possède la propriété de Feller.

Théorème 3.5.3. *Le semigroupe de Feller P_t correspondant à la solution généralisée de l'équation (3.27) admet une mesure invariante.*

Ce résultat de Flandoli [24] repose sur le caractère dissipatif et des majorations *a priori* sur l'équation

$$dV_a(t) + AV_a(t)dt + B\left(V_a(t)+W_A^a(t), V_a(t)+W_A^a(t)\right)dt = \left(aW_A^a(t)+f\right)dt$$

vérifiée par $V_a := X - W_A^a$ où W_A^a est la convolution stochastique modifiée définie par (3.25). Ces majorations lui permettent en fait d'établir le

Lemme 3.5.4. *[24] Pour $t_0 \in \mathbb{R}$ et $t \ge t_0$, on note $X(t,t_0)$ la solution de (3.27) au temps t partie de $x=0$ en t_0. Il existe $r(\omega)$, une variable aléatoire réelle $\mathbb{P}$-p.s finie telle que*

$$\sup_{-\infty<t_0\le 0} |A^{\frac{1}{2}}X(0,t_0)| \le r.$$

En particulier, la famille de variables aléatoires $\{X(0,t_0) \ \ , \ -\infty < t_0 \leq 0\}$ est bornée en probabilité dans $D(A^{\frac{1}{2}})$. Pour $0 \leq \alpha < \frac{1}{2}$ l'injection $D(A^{\frac{1}{2}}) \subset D(A^{\alpha})$ est compacte, donc la famille de lois correspondante est tendue dans $D(A^{\alpha})$. Comme $\mathcal{L}X(t,t_0) = P^*_{|t_0|}\delta_0$, on en déduit que la famille de loi $\left(1/T \int_0^T P_t^* \delta_0 dt\right)_{T\geq 0}$ est tendue. D'après le critère de Krylov-Bogoliubov (3.2), le semigroupe P_t admet une mesure invariante dans $D(A^{\alpha})$ (et donc dans H).

Nous présentons maintenant les deux résultats cruciaux dus à Ferrario [22], donnant les conditions (3.6) requises pour l'étude des grandes déviations.

Théorème 3.5.5. *Soit $0 < \alpha < \frac{1}{2}$, et un opérateur G vérifiant l'hypothèse (3.28). Si de plus $Im(G)$ est dense dans $D(A^{\frac{1}{2}})$ pour la topologie du graphe de $A^{\frac{1}{2}}$, alors le semigroupe P_t associé à la solution de (3.27) est topologiquement irréductible sur $D(A^{\alpha})$.*

Démonstration. On donne quelques étapes de la preuve de ce résultat. Il s'agit de montrer que l'on a $\mathbb{P}\left(|X(t,x) - y|_{\alpha} < \varepsilon\right) > 0$, pour tout $t > 0$, tout $x, y \in D(A^{\alpha})$ et tout $\varepsilon > 0$. Dans [22], Ferrario considère l'application

$$\Phi : z \in C_0\left([0,t], D(A^{\alpha})\right) \cap L^{\frac{4}{1-2\alpha}}\left(0,t; D(A^{\frac{1}{4}+\frac{\alpha}{2}})\right) \quad \rightarrow \quad v + z \in C\left([0,t], D(A^{\alpha})\right)$$

où v est la solution de l'équation (3.29), avec la force auxiliaire z à la place de W_A, de sorte que $\Phi(W_A) = X$ (Théorème 3.5.2).

Elle détaille ensuite la construction d'une courbe $\bar{u} \in C([0,t], D(A^{\alpha}))$ reliant x à y, i.e vérifiant $\bar{u}(0) = x$ et $\bar{u}(t) = y$, ainsi que celle d'une courbe $\bar{z}$ telle que $\phi(\bar{z}) = \bar{u}$ et possédant la régularité $\bar{z} \in C_0([0,t], D(A^{\alpha})) \cap L^{\frac{4}{1-2\alpha}}\left(0,T; D(A^{\frac{1}{4}+\frac{\alpha}{2}})\right)$.

Puis, elle montre que Φ est Lipschitz pour les topologies correspondantes, de sorte que l'on ait le contrôle

$$\|X - \bar{u}\|_{C([0,t],D(A^{\alpha}))} \quad \leq \quad C\left(\|W_A - \bar{z}\|_{C([0,t],D(A^{\alpha}))} + \|W_A - \bar{z}\|_{L^{\frac{4}{1-2\alpha}}\left(0,t;D(A^{\frac{1}{4}+\frac{\alpha}{2}})\right)}\right)$$

Pour conclure, on remarque enfin que

$$\begin{aligned}
& \mathbb{P}\left(|X(t,x) - y|_{\alpha} < \varepsilon\right) \\
\geq \ & \mathbb{P}\left(\|X - \bar{u}\|_{C([0,t],D(A^{\alpha}))} < \varepsilon\right) \\
\geq \ & \mathbb{P}\left(\|W_A - \bar{z}\|_{C([0,t],D(A^{\alpha}))} + \|W_A - \bar{z}\|_{L^{\frac{4}{1-2\alpha}}\left(0,t;D\left(A^{\frac{1}{4}+\frac{\alpha}{2}}\right)\right)} < \varepsilon/C\right) \\
> \ & 0
\end{aligned}$$

puisque la loi de W_A est de mesure pleine dans $C([0,t], D(A^{\frac{1}{2}}))$ grâce à l'hypothèse de densité de $Im(G)$ (Proposition 3.4.2), donc aussi dans les deux espaces ci-dessus comme $\frac{1}{4} + \frac{\alpha}{2} < \frac{1}{2}$ pour $\alpha < \frac{1}{2}$. $\square$

On renforce maintenant l'hypothèse (3.28) sur le bruit par une condition de non dégénérescence. On suppose que l'opérateur G est injectif et vérifie pour un certain $\frac{1}{4} < \alpha < \frac{1}{2}$ et un $0 < \varepsilon \le 2\alpha - \frac{1}{2}$,

$$D(A^{2\alpha}) \subset Im(G) \subset D(A^{\frac{1}{2}+\varepsilon}). \tag{3.31}$$

La condition $\alpha > \frac{1}{4}$ est une condition de compatibilité avec la seconde inclusion. On impose aussi $\alpha < \frac{1}{2}$ pour rester dans le cadre de la partie (ii) du théorème 3.5.2 (voir la section suivante pour l'interprétation physique), pour avoir l'irréductibilité sur $D(A^{2\alpha})$, et enfin pour la majoration (3.36) ci dessous. C'est suffisant pour établir le

Théorème 3.5.6. *Soit $0 < \alpha < \frac{1}{2}$, et G un opérateur vérifiant l'hypothèse (3.31). Le semigroupe P_t associé à la solution de (3.27) est fortement Feller sur $D(A^\alpha)$.*

Démonstration. On souligne le rôle de la condition de non dégénérescence sur le bruit en donnant quelques étapes de cette preuve détaillée dans le travail de Ferrario [22].

Il s'agit de montrer que pour tout $t > 0$, et toute fonction $\Psi \in b\mathcal{B}(D(A^\alpha))$,

$$|P_t\Psi(x) - P_t\Psi(y)| \to 0 \quad \text{si} \quad |x-y|_\alpha \to 0 \tag{3.32}$$

De façon usuelle, on tronque d'abord la non linéarité de l'équation (3.27) par une fonction plateau C^∞, notée Θ_R, prenant la valeur 1 sur $[-R, R]$, et 0 en dehors de $[-R-1, R+1]$. On regarde ainsi

$$dX^{(R)}(t) + AX^{(R)}(t)dt + \Theta_R(|A^\alpha X^{(R)}(t)|)\ B\left(X^{(R)}(t), X^{(R)}(t)\right) dt = f dt + GdW(t)$$

avec la condition initiale $X^{(R)}(0) = x$, ainsi que $P_t^{(R)}$ le semigroupe associé. Les processus X et $X^{(R)}$ coïncident tant que X est dans la boule

$$B_R := \{z \in D(A^\alpha)\ :\ |z|_\alpha \le r\}.$$

Avec des estimations a priori dans $C([0,T], D(A^\alpha))$ qui sont celles utilisées pour prouver l'existence de solutions, on montre que l'on a $|P_t\Psi(x) - P_t^{(R)}\Psi(x)| \to 0$ lorsque $R \to \infty$ uniformément sur toutes les boules $B(x,r)$ de $D(A^\alpha)$. Ainsi, prouver (3.32) revient à établir que $P_t^{(R)}$ est fortement Feller.

Pour ce faire, Ferrario montre d'abord que le semigroupe du processus de Markov noté $X_n^{(R)}$ sur $H_n := Vect(e_1, \dots, e_n)$ solution du système d'approximations finidimensionnelles est Lipschitz Feller avant de passer à la limite. Plus précisément, il faut vérifier que pour tout $t, R > 0$,

$$|P_{t,n}^{(R)}\Psi(x) - P_{t,n}^{(R)}\Psi(y)| \le L(R,t)|x-y|_\alpha \tag{3.33}$$

où $L(R,t)$ est une constante indépendante de n et $P_{t,n}^{(R)}$ est le semigroupe du processus $X_n^{(R)}$ solution de

$$dX_n^{(R)}(t) + AX_n^{(R)}(t)dt + \Theta_R(|A^\alpha X_n^{(R)}(t)|)\ B\left(X_n^{(R)}(t), X_n^{(R)}(t)\right) dt$$

$$= \Pi_n f dt + \Pi_n G dW(t)$$

avec la condition initiale $X_n^{(R)}(0) = \Pi_n x$, et $\Pi_n : H \to H_n$ la projection orthogonale.

Dans ce cas, la formule de la moyenne s'écrit

$$|P_{t,n}^{(R)}\Psi(x) - P_{t,n}^{(R)}\Psi(y)| \leq \sup_{k,h\in H_n\ ,\ |h|_\alpha\leq 1} |DP_{t,n}^{(R)}\Psi(k)\cdot h|\ |x-y|_\alpha \tag{3.34}$$

où $DP_{t,n}^{(R)}\Psi(k)\cdot h \in H_n$ est la dérivée de l'application $y \in D(A^\alpha) \to P_{t,n}^{(R)}\Psi(y) \in \mathbb{R}$ au point k, dans la direction h. La formule de Bismut-Elworthy [20] est systématiquement utilisée [22, 23, 25] pour estimer la dérivée précédente : pour $\beta^n := (\beta_1, \ldots, \beta_n)$ un mouvement Brownien de dimension n, on a

$$\begin{aligned}
|DP_{t,n}^{(R)}\Psi(k)\cdot h| &= \left|\frac{1}{t}\mathbb{E}\left[\Psi(X_n^{(R)}(t,k))\int_0^t \langle(\Pi_n GG^*\Pi_n)^{-\frac12}\ DX_n^{(R)}(s,k)\cdot h\ ,\ d\beta^n(s)\rangle\right]\right| \\
&\leq \frac{1}{t}\|\Psi\|_\infty\left[\mathbb{E}\left|\int_0^t \langle(\Pi_n GG^*\Pi_n)^{-\frac12}\ DX_n^{(R)}(s,k)\cdot h\ ,\ d\beta^n(s)\rangle\right|^2\right]^{\frac12} \\
&\leq \frac{1}{t}\|\Psi\|_\infty\left[\mathbb{E}\int_0^t |(\Pi_n GG^*\Pi_n)^{-\frac12}\ DX_n^{(R)}(s,k)\cdot h|^2 ds\right]^{\frac12}
\end{aligned}$$

C'est ici que la condition de non dégénérescence (3.31) est cruciale. Elle nous permet de prouver

$$|(\Pi_n GG^*\Pi_n)^{-\frac12}y|^2 \leq C|A^{2\alpha}y|^2, \quad \forall y \in H_n \tag{3.35}$$

pour une constante $C > 0$ indépendante de de n et de y. En effet, cette condition signifie que l'opérateur $G^{-1}A^{-2\alpha}$ est borné sur H, de même que son adjoint. En particulier, on obtient pour $x \in H$

$$\begin{aligned}
\langle(A^{2\alpha}GG^*A^{2\alpha})^{-1}x, x\rangle &= \langle(G^{-1}A^{-2\alpha})^*G^{-1}A^{-2\alpha}x, x\rangle \\
&= |G^{-1}A^{-2\alpha}x|^2 \\
&\leq C|x|^2.
\end{aligned}$$

Comme Π_n et A commutent, l'inégalité précédente implique que pour tout $x \in H_n$,

$$\langle(A^{2\alpha}\Pi_n GG^*\Pi_n A^{2\alpha})^{-1}x, x\rangle = \langle(A^{2\alpha}GG^*A^{2\alpha})^{-1}x, x\rangle \leq C|x|^2.$$

Enfin, $A^{2\alpha}$ étant un isomorphisme sur H_n, on en déduit que pour tout $y \in H_n$,

$$\begin{aligned}
\left|(\Pi_n GG^*\Pi_n)^{-\frac12}y\right|^2 &= \langle(\Pi_n GG^*\Pi_n)^{-1}y, y\rangle \\
&\leq C|A^{2\alpha}y|^2,
\end{aligned}$$

ce qui implique (3.35). En appliquant ce contrôle à $y = DX_n^{(R)}(s,k)\cdot h$, on obtient

$$|DP_{t,n}^{(R)}\Psi(k)\cdot h| \;\le\; \frac{C}{t}\|\Psi\|_\infty \left[\mathbb{E}\int_0^t |A^{2\alpha}\ DX_n^{(R)}(s,k)\cdot h|^2 ds\right]^{\frac{1}{2}}.$$

L'inégalité suivante (sans restriction sur $\alpha \ge 1/4$), déja obtenue par Flandoli et Maslowski [25], repose sur des estimations a priori de $DX_n^{(R)}(s,k)\cdot h$ assez classiques

$$C\mathbb{E}\int_0^t |A^{2\alpha}\ DX_n^{(R)}(s,k)\cdot h|^2 ds \le C_{1,R,t}^2 |A^{2\alpha-\frac{1}{2}}h|^2.$$

Enfin, dès que $\alpha \le 1/2$, on a aussi

$$C_{1,R,t}^2|A^{2\alpha-\frac{1}{2}}h|^2 \le C_{2,R,t}^2|A^\alpha h|^2. \tag{3.36}$$

Les constantes $C_{i,R,t}$ sont ici indépendantes de n et de $h,k \in H_n$, et on obtient

$$|DP_{t,n}^{(R)}\Psi(k)\cdot h| \;\le\; \frac{C_{2,R,t}}{t}\|\Psi\|_\infty\ |h|_\alpha.$$

Finalement, dans l'expression (3.34), on a

$$\sup_{k,h\in H_n\ ,\ |h|_\alpha\le 1} |DP_{t,n}^{(R)}\Psi(k)\cdot h| \le \frac{C_{2,R,t}}{t}\|\Psi\|_\infty$$

ce qui permet d'obtenir (3.33).

□

Comme la condition (3.31) entraîne en particulier que $Im(G)$ est dense dans $D(A^{\frac{1}{2}})$, on déduit des deux résultats précédents le

Corollaire 3.5.7. *Soit $0 < \alpha < \frac{1}{2}$, et un opérateur G vérifiant l'hypothèse (3.31). Le semigroupe P_t associé à la solution de (3.27) admet une unique mesure invariante sur $D(A^\alpha)$.*

3.6 Présentation des résultats du chapitre 4

3.6.1 A propos du bruit

On suppose que le bruit satisfait à la condition (3.31) pour un $\alpha \in (1/4, 1/2)$ et un $\varepsilon \in (0, 1-2\alpha)$. Par exemple, on remarque que $G = A^{-\beta}L$ vérifie cette hypothèse dès que L est un opérateur linéaire inversible de H et $\beta \in (1/2, 2\alpha]$. Le cas le plus simple et le plus courant survient lorsque $Ge_k = \sigma_k e_k$, le bruit prenant alors la forme

$$GW(t) = \sum_{k\ge 1}\sigma_k\beta_k(t)e_k. \tag{3.37}$$

Connaissant le comportement (3.17) du spectre de A, la condition (3.31) est alors équivalente à

$$\frac{c}{k^{2\alpha}} \leq \sigma_k \leq \frac{C}{k^{\frac{1}{2}+\varepsilon}}$$

pour un $\alpha \in (1/4, 1/2)$, un $\varepsilon \in (0, 1-2\alpha)$ et deux constantes strictement positives c et C.

Physiquement, la première inégalité est la condition de non dégénerescence qui assure que l'ensemble des modes est excité avec une amplitude suffisante pour le mélange du système.

La seconde inégalité, est une condition naturelle dans le cas d'une étude de stabilité : elle garantit un bruit de classe trace, ce qui signifie que l'énergie injectée par unité de temps est finie. En effet, on a

$$\mathbb{E}\left|\int_0^t GdW(s)\right|^2 = t\ tr(Q) < \infty$$

où une petite équation aux dimensions (notées $[\ \cdot\ ]$) permet de se convaincre que $t\ tr(Q)$ est homogène à une énergie : notant L, T, E les dimensions respectives de distance, temps et énergie on a (en oubliant les densités) $E = L^2/T^2$. D'après l'équation on a $[dX(t)] = [Ge_k d\beta_k(t)] = L/T$. Comme $[d\beta_k(t)]^2 = [dt] = T$, on voit que $[Ge_k] = L/T^{\frac{3}{2}}$ et donc $[tr(Q)] = [Ge_k]^2 = L^2/T^3 = E/T$.

3.6.2 Plan de l'étude

On commence par établir des estimations exponentielles de la solution. En travaillant sur des approximations fini-dimensionnelles avec la formule de Itô, la condition de trace finie du bruit permet d'établir la propriété suivante. On note encore X la solution de l'équation (3.27) et $\|Q\| := \|Q\|_{L(H)}$ la norme d'opérateur de Q.

Proposition 3.6.1. *Soit* $0 < \lambda_0 < \frac{\lambda_1}{2\|Q\|}$, *on a pour tout* $x \in H$,

$$\mathbb{E}^x \exp\left(\lambda_0 |X(t)|^2 ds\right) \leq \exp\left(\lambda_0 t \left(tr(Q) + \frac{|f|^2}{\lambda_1 - 2\|Q\|\lambda_0}\right)\right) e^{\lambda_0 |x|^2}$$

et

$$\mathbb{E}^x \exp\left(\lambda_0 \int_0^t |A^{\frac{1}{2}} X(s)|^2 ds\right) \leq \exp\left(\lambda_0 t \left(tr(Q) + \frac{|f|^2}{\lambda_1 - 2\|Q\|\lambda_0}\right)\right) e^{\lambda_0 |x|^2}. \quad (3.38)$$

Ces estimations exponentielles sont centrales. Elle permettent par exemple de prouver que l'entropie de Donsker-Varadhan $J : \nu \in M_1(H) \to J(\nu) \in [0, +\infty]$ définie par (3.10) vérifie pour tout $0 < \lambda_0 < \frac{\lambda_1}{2\|Q\|}$

$$\lambda_0 \int |A^{\frac{1}{2}} x|^2 d\nu \leq J(\nu) + \lambda_0 \left(tr(Q) + \frac{|f|^2}{\lambda_1 - 2\|Q\|\lambda_0}\right).$$

On montre ainsi que J satisfait le

Lemme 3.6.2. *Soit μ l'unique mesure invariante de l'équation (3.27). On a $[J=0]=\{\mu\}$. De plus, $\forall \nu \in M_1(H)$,*

$$J(\nu) < +\infty \implies \nu \ll \mu,\ \nu(V)=1 \text{ et } \int_V |A^{\frac{1}{2}}x|^2 d\nu < +\infty. \tag{3.39}$$

Dans le chapitre suivant, on donne aussi une preuve alternative (basée sur la théorie des grandes déviations) de l'existence d'une mesure invariante pour un processus Feller-Markov $X(t)$ qui vérifie une estimation exponentielle du type (3.38).

Pour obtenir le PGD de la définition 3.2.4, dans un premier temps, on utilise les Théorèmes 3.2.5, 3.5.5 et 3.5.6 pour l'établir sur l'espace $E = D(A^{\alpha})$. Pour cela, il faut exhiber des compacts de l'espace $D(A^{\alpha})$. Comme on a $\alpha < 1/2$, l'injection de $D\left(A^{\frac{1}{2}}\right)$ dans $D(A^{\alpha})$ est compacte, et les boules de $D\left(A^{\frac{1}{2}}\right)$ deviennent ainsi des candidats raisonnables. De plus, à partir des estimations exponentielles précédentes, on peut définir un ensemble $\mathcal{M}^*_{\lambda_0,L}$ de mesures initiales sur $D\left(A^{\alpha}\right)$, et trouver un compact vérifiant les conditions de retour (3.7) du critère. On établit plus précisément le

Lemme 3.6.3. *Pour $0 < \lambda_0 < \frac{\lambda_1}{2\|Q\|}$ et $L \geq 1$, on considère l'ensemble*

$$\mathcal{M}^*_{\lambda_0,L} := \left\{\nu \in M_1\left(D\left(A^{\alpha}\right)\right)\ /\ \int \exp\left(\lambda_0|x|^2\right)\nu(dx) \leq L\right\}. \tag{3.40}$$

On a alors pour toute mesure $\nu \in \mathcal{M}^*_{\lambda_0,L}$

$$\mathbb{E}^{\nu} \exp\left(\lambda_0 \int_0^t |A^{\frac{1}{2}}X(s)|^2 ds\right) \leq \exp\left(\lambda_0 t\left(tr(Q) + \frac{|f|^2}{\lambda_1 - 2\|Q\|\lambda_0}\right)\right) L. \tag{3.41}$$

Comme $\alpha < \frac{1}{2}$, l'ensemble

$$K := \left\{x \in D(A^{\frac{1}{2}})\ t.q\ |A^{\frac{1}{2}}x| \leq M\right\}$$

est compact dans $D\left(A^{\alpha}\right)$. De plus, pour chaque $\lambda > 0$, on peut choisir M suffisamment grand pour que les conditions du Théorème 3.2.5 soient vérifiées, par K et $\mathcal{M} := \mathcal{M}^*_{\lambda_0,L}$.

On en déduit que la famille $\mathbb{P}_\nu(L_T \in \cdot)$ pour $T \to +\infty$ vérifie le PGD de la Définition 3.2.4 sur $M_1(D\left(A^{\alpha}\right))$, uniformément pour les mesures initiales $\nu \in \mathcal{M}^*_{\lambda_0,L}$.

Une seconde difficulté est d'étendre ce PGD pour des conditions initiales de l'espace H, des ouverts et des fermés de $M_1(H)$. A l'aide des propriétés de régularité de la solution, comme le (iii) du Théorème 3.5.2, et en utilisant le caractère locale d'une borne inférieure ainsi que la forme générale des voisinages pour la topologie τ, on prouve le

Théorème 3.6.4. *Soit $L > 1$, $0 < \lambda_0 < \frac{\lambda_1}{2\|Q\|}$ et une mesure $\nu \in \mathcal{M}_{\lambda_0,L}$, où*

$$\mathcal{M}_{\lambda_0,L} := \left\{ \nu \in M_1(H) \ / \ \int \exp \lambda_0 |x|^2 \nu(dx) \leq L \right\}$$

alors la loi $\tilde{\nu} := \mathbb{P}_\nu (X(1) \in dy)$ vérifie
(a) $\tilde{\nu}(D(A^\alpha)) = 1$
(b)$\int_H e^{\lambda_0|x|^2} \tilde{\nu}(dx) \leq e^{\lambda_0 C} L$, où $C = tr(Q) + \frac{|f|^2}{\lambda_1 - 2\|Q\|\lambda_0}$
En particulier $\tilde{\nu} \in \mathcal{M}^*_{\lambda_0, e^{\lambda_0 C} L}$, et $\mathbb{P}_\nu(L_T \circ \theta_1 \in \cdot) = \mathbb{P}_{\tilde{\nu}}(L_T \in \cdot)$ pour $\theta_1 X_t = X_{t+1}$.

On déduit de cela que la famille $\mathbb{P}_\nu(L_T \in \cdot)$ pour $T \to +\infty$ vérifie le Principe de Grandes Déviations de la Définition 3.2.4 sur $M_1(H)$, uniformément pour les mesures initiales $\nu \in \mathcal{M} := \mathcal{M}_{\lambda_0,L}$.

On peut noter que la classe $\mathcal{M}_{\lambda_0,L}$ des distributions initiales est suffisamment riche : elle contient les mesures de dirac δ_x avec x dans n'importe quelle boule de H centrée en 0, quitte à choisir L assez grand.

Il faut aussi souligner qu'un PGD avec la topologie τ est plus fort qu'avec la topologie usuelle w de la convergence étroite [18]. Dans le cas des équations de Navier-Stokes, la topologie τ permet par exemple de déduire des PGD sur d'autres quantités. Plus précisément, on établit le

Théorème 3.6.5. *Soit $(\mathbb{B}\ ,\ \|\cdot\|_{\mathbb{B}})$ un espace de Banach séparable, et soit g une fonction mesurable $g : D\left(A^{\frac{1}{2}}\right) \to \mathbb{B}$ bornée sur les boules du type $\{x \ \ t.q \ \ |A^{\frac{1}{2}}x| \leq R\}$ pour tout $R > 0$, et telle que*

$$\lim_{|A^{\frac{1}{2}}x| \to \infty} \frac{\|g(x)\|_{\mathbb{B}}}{|A^{\frac{1}{2}}x|^2} = 0.$$

Alors la famille $\mathbb{P}_\nu (L_T(g) \in \cdot)$ vérifie un PGD sur $\mathbb{B}$ de fonction de taux I_g donnée par

$$I_g(y) = \inf\{J(\nu); \quad J(\nu) < +\infty, \ \nu(g) = y\}, \ \forall y \in \mathbb{B}$$

uniformément pour les conditions initiales ν in $\mathcal{M}_{\lambda_0,L}$ (pour tout $L > 1$).

Ce théorème s'applique par exemple au cas $\mathbb{B} = \mathbb{R}$ et $g(x) = |A^{\frac{1}{2}}x|$, ce qui donne un PGD sur $\mathbb{R}$ pour la famille

$$\mathbb{P}_\nu \left(\frac{1}{T} \int_0^T |X(t)|_V dt \ \in \ \cdot \right), \quad T > 0$$

où la quantité $|X(t)|_V$ présente un intérêt physique certain : dans le cas d'un fluide incompressible, avec des conditions périodiques ou nulles aux bords, elle est égale à l'amplitude de la vitesse angulaire locale du fluide (voir [11] p17).

Comme on l'a dit, la description de la vitesse exacte de convergence exponentielle donnée par notre PGD est surtout qualitative. On propose cependant dans le chapitre 4 un exemple où J est donnée de façon un peu plus explicite, au moins pour certaines mesures. Si $h \in C_1(H)$ vérifie $\langle GG^* \nabla^H h(x), \nabla^H h(x)\rangle \leq C < \infty$ où ∇^H est le gradient dans H, et pour G symétrique, on note μ^h l'unique mesure invariante du semigroupe associé à l'équation perturbée

$$dX_t + AX_t dt + B(X_t, X_t)\, dt = f dt + \sqrt{GG^*}\nabla^H h(X_t)\, dt + G d\tilde{W}_t \quad , \quad X_0 = x,$$

semigroupe défini à l'aide du théorème de Girsanov. On a alors

$$J(\mu^h) \;=\; \frac{1}{2}\int_H \langle GG^* \nabla^H h, \nabla^H h\rangle d\mu^h,$$

Par contre, on ne peut estimer $J(\nu)$ pour des mesures ν plus générales : l'inégalité de Sobolev logarithmique pour l'équation NS2D semble pour l'instant inconnue (voir cependant les récents travaux de Goldys-Maslowski [28] et Hairer-Mattingly [30] pour l'existence d'un trou spectral dans des espaces pondérés de fonctions bornées ou Lipschitz).

3.6.3 Quelques remarques et perspectives...

Pour clore cette section, nous présentons quelques développements très récents à propos du comportement ergodique des équations de Navier-Stokes stochastiques.

Le cas des conditions périodiques

On peut conduire la même étude si on remplace les conditions de type Dirichlet par des conditions périodiques. Dans le cas du Tore $\mathbb{T}^2 := [0, 2\Pi]^2$, par exemple, la condition (3.12) est remplacée par une condition de champs moyen nul

$$\int_{\mathbb{T}^2} u(0,\xi) d\xi \;=\; \int_{\mathbb{T}^2} u(t,\xi) d\xi \;=\; 0, \qquad \forall t \geq 0$$

et la condition périodique aux bords

$$u(t, \xi + 2\pi\eta_i) = u(t,\xi), \qquad \forall t \geq 0, \xi \in \mathbb{R}^2, i = 1,2$$

où $\{\eta_1, \eta_2\}$ est la base canonique de $\mathbb{R}^2$. Ce type de condition a moins de sens physique, mais a l'avantage de simplifier le cadre fonctionnel pour l'étude de l'équation. Par exemple, on a alors $A = \Delta$. Ainsi les valeurs propres λ_k, et les éléments e_k de la base peuvent être explicités : on rentre dans le cadre des séries de Fourier. De plus, la forme bilinéaire B satisfait à quelques propriétés supplémentaires intéressantes, comme l'orthogonalité $\langle B(u,v), Av\rangle = 0$ (en dimension 2).

Dans un travail récent, Ferrario [23] a prouvé dans ce nouveau cadre, pour tout $r > 0$ fixé, la propriété forte de Feller et l'irréductibilité topologique dans $D(A^r)$ sous la condition $D(A^{\frac{1}{2}+r}) \subset Im(G) \subset D(A^{r+\varepsilon})$ Pour r grand, cette étude permet de diminuer largement l'hypothèse de non dégénerescence.

Pour les grandes déviations, on doit garder la condition $Im(G) \subset D(A^{1/2+\varepsilon})$, pour que l'opérateur $Q = GG^*$ soit de trace finie, et que la proposition 3.6.1 reste valide. Cela permet d'obtenir les conditions (3.7) pour des compacts de $D(A^r)$ avec $r < 1/2$, et de montrer le PGD de la définition 3.2.4 sur H sous une condition semblable à (3.31) qui s'écrit $D(A^{\frac{1}{2}+r}) \subset Im(G) \subset D(A^{1/2+\varepsilon})$ pour $r \in (0, 1/2)$ et $\varepsilon \in (0, r)$.

Cependant grâce à la propriété d'orthogonalité de B, on peut améliorer la proposition 3.6.1 en s'assurant que $tr(GAG^*) < \infty$, par exemple si $Im(G) \subset D(A^{1+\varepsilon})$. Cela permet d'obtenir les conditions (3.7) pour des compacts de $D(A^r)$ jusqu'à $r < 1$, et donc le PGD sous la condition

$$D(A^{\frac{1}{2}+r}) \subset Im(G) \subset D(A^{1+\varepsilon})$$

pour $r \in (1/2, 1)$.

Le cas des bruits dégénérés

Dans le cas de bruits dégénérés, c'est à dire lorsque $\sigma_k = 0$ sauf pour un nombre fini d'indices k dans (3.37), la situation est plus délicate. Le semigroupe P_t ne vérifie plus la propriété forte de Feller. Cependant, Bricmont, Kupiainen, Lefevere [3] et Weinan E, Mattingly, Sinai [21] ont obtenu l'unicité de la mesure invariante, ainsi que la convergence exponentielle si les σ_k sont non nuls quand $1 \leq k \leq N_\nu$, pour un certain nombre minimum de modes excités N_ν qui se comporte en ν^{-3} lorsque $\nu \to 0$. Ils ont utilisé une méthode introduite par Kuksin et Shirikyan [32] et réduisent l'étude à celle des bas modes, les hauts modes étant représentés comme des fonctions de tout le passé des bas modes : la dynamique n'est plus Markovienne.

Plus récemment, Hairer, Mattingly et Pardoux [37, 29] ont prouvé l'ergodicité sans la dépendance en la viscosité, et pour une condition purement géométrique sur le nombre N de modes à forcer. Dans ce cadre hypoelliptique, ils ont utilisé des outils de calcul de Malliavin pour montrer que le semigroupe P_t vérifie une propriété *asymptotiquement forte de Feller* du type

$$|\nabla P_t \phi(x)| \leq C(|x|) \ \ (\|\phi\|_\infty + \delta_t \|\nabla \phi\|_\infty)$$

où $\delta_t \to 0$ lorsque $t \to \infty$. Ce résultat ainsi qu'une propriété d'irréductibilité leur permet d'établir l'unicité de la mesure invariante.

Sans la propriété forte de Feller, le PGD pour la topologie τ ne peut être prouvé et semble faux. La norme en variation totale n'est plus adaptée. Cependant, un trou spectral du semigroupe a pu être établi dernièrement par Hairer et Mattingly [30] pour une distance de Wasserstein. Le Principe de Grandes Déviations pour la topologie faible w est un problème ouvert intéressant.

Le cas 3-D

Le cas tridimensionnel est beaucoup plus délicat. L'existence d'une solution faible globale à l'équation peut s'obtenir par des techniques analogues, dans les cas déterministes et stochastiques. Par contre, le fameux problème de l'unicité empêche d'utiliser les arguments précédents pour définir un semigroupe de transition. Il faut nécessairement considérer des solutions faibles au sens probabiliste : l'espace de probabilité de base n'est pas fixé, son existence est affirmée dans la définition du concept de solution. Le problème ouvert de l'unicité en loi qui se pose alors semble plus abordable.

En considérant l'équation de Kolmogorov associée, Da Prato et Debussche [8] ont montré l'existence d'un semigroupe de Markov, limite d'approximations de Galerkin, associé aux solutions faibles de l'équation. Des travaux de Flandoli et Romito [26], puis Debussche et Odasso [14] affinent cette procédure. Dans [14], des bruits multiplicatifs sont aussi traités, et sous des conditions de non dégénérescence, le semigroupe obtenu vérifie les propriétés fortes de Feller et d'irréductiblité topologique.

Ces différents résultats devraient permettre d'établir le PGD du critère 3.2.5. Enfin, il faut noter qu'un résultat de convergence exponentielle a été prouvé dans le cadre précédent par Odasso [40] en utilisant des arguments de couplage.

3.7 Présentation des résultats du chapitre 5

Prenons le cas de conditions aux bords périodiques pour compléter les équations de Navier-Stokes. Le rôle central est joué par l'espace $\mathbb{L}^2$ fermeture pour la topologie L^2 de l'ensemble des champs de vecteurs (u_1, u_2) qui soient C^∞ à divergence nulle et de moyenne nulle sur $\mathbb{T}^2$. On note (e_j) la base de Fourier de $\mathbb{L}^2$. La formulation abstraite (3.13) s'écrit alors

$$du(t,\xi) + \nu Au(t,\xi)dt + B(u,u)dt = f(t,\xi)dt, \quad t \geq 0, \xi \in \mathbb{T}^2 \tag{3.42}$$

où $\mathbb{T}^2 = [0, 2\pi]^2$ est le tore.

Plutôt que de considérer un bruit blanc en temps, certains auteurs ont suggéré l'ajout d'une δ-force dite perturbation sous forme de *kicks*. La force considérée s'écrit alors

$$f(t,\xi) = \sum_{k\geq 0} \delta(t-k)\eta_k(\xi) \tag{3.43}$$

où les η_k sont des v.a.i.i.d dans $\mathbb{L}^2$. Cela signifie que la force change $u(k,\xi)$ en $u(k+0,\xi) = u(k,\xi)+\eta_k$ alors que, pour $k \leq t < k+1$, $u(t,\xi)$ vérifie l'équation (3.42) avec $f = 0$ et la condition initiale $u(k+0,\xi)$. Plus précisément, pour $\{S_t,\ t \geq 0\}$ le flow de l'équation (3.42) avec $f = 0$, les valeurs de la solution u sur $[k, k+1)$ sont données pour tout entier k par

$$u(t) := \begin{cases} S_{t-k}u(k+0), \text{ pour } k \leq t < k+1 \\ S_1 u(k+0) + \eta_{k+1} \text{ pour } t = k+1 \end{cases}$$

Ainsi, l'étude du comportement en temps grand de l'équation perturbée par la force (3.43) revient à l'étude du système dynamique

$$u_k = S(u_{k-1}) + \eta_k \tag{3.44}$$

où $S = S_1$ et $u_k = u(k+0,\cdot) \in \mathbb{L}^2$.

Pour obtenir des résultats précis sur le comportement ergodique, il faut naturellement faire des hypotèses sur la structure du bruit. Ce modèle a été introduit par Kuksin et Shirikyan [32, 33] qui ont d'abord regardé l'unicité de la mesure invariante pour des v.a.i.i.d η_k de la forme

$$\eta_k = \sum_{j \geq 1} b_j \xi_{j,k} e_j$$

où la distribution de la variable aléatoire réelle $\xi_{j,k}$ est indépendante de k et possède une densité par rapport à la mesure de Lebesgue avec de bonnes propriétés, tandis que les b_j sont des coefficient réels . Ils ont établi l'ergodicité sous une condition de non dégénérescence i.e $b_j \neq 0$ pour les premiers rangs. Ils ont ensuite prouvé dans [32] (voir aussi Masmoudi et Young [35]) la convergence exponentielle vers la solution stationnaire lorsque les $\xi_{j,k}$ sont à support borné. Enfin, Shirikyan [41] a récemment étendu ce résultat au cas non borné, lorsque la densité des $\xi_{j,k}$ admet certains moments exponentiels. Pour ces différents exemples, où un bon comportement ergodique est prouvé, il serait intéressant d'établir un principe de grandes déviations.

Dans ce chapitre, on laisse de côté le problème de l'unicité de la mesure d'équilibre et du comportement asymptotique. On cherche des conditions simples sur η_k permettant d'assurer l'existence d'une mesure invariante. On démontre en effet, le

Théorème 3.7.1. *Si il existe $c_0 > 0$ tel que*

$$\mathbb{E} \log(1 + c_0 \|\eta_k\|_{\mathbb{L}^2}) < \infty \tag{3.45}$$

alors la chaîne de Markov (3.44) *associée aux équations de Navier-Stokes admet une mesure invariante.*

En fait, on établit un critère d'existence de mesure invariante dans un cadre plus général. Soit $(E, |\cdot|)$ un espace de Banach, $B(R)$ la boule centrée en 0 de rayon R. On montre d'abord le

Théorème 3.7.2. *Soit P un noyau de transition de Feller. On suppose que*
(i) il existe une fonction $u : E \to [0, \infty]$ bornée sur les boules telle que

$$\exists R, b > 0, \quad \forall x \in E, \quad Pu(x) \leq u(x) - 1 + b\mathbb{1}_{B(R)}(x),$$

(ii) pour tout $R > 0$, la famille de lois $\{P(x, dy),\ x \in B(R)\}$ est tendue.
Alors P admet une mesure invariante.

Les conditions sur la dérive de type (i) pour les chaînes de Markov ont été étudiées par Meyn et Tweedie [39] lorsque $B(R)$ est remplacée par un compact. Le résultat que l'on obtient est plus facile à manipuler, en particulier dans le cas d'un espace E de dimension infinie. On en déduit le

Théorème 3.7.3. *On considère le système dynamique dans E défini par*

$$X_{k+1} := S(X_k) + W_{k+1} \tag{3.46}$$

où $(W_k)_{k\geq 1}$ est une suite de v.a.i.i.d dans E. On suppose que
(a) l'opérateur $S : E \to E$ est continu et pour tout $R > 0$, l'ensemble $S\left(B(R)\right)$ est relativement compact
(b) il existe $0 < r < 1$ et $R_0 > 0$ tels que pour tout $|x| \geq R_0$, $\quad |S(x)| \leq r|x|$
(c) il existe $c_0 > 0$ tel que $\mathbb{E}\log(1 + c_0|W|) < \infty$.
Alors le système (3.46) possède une mesure invariante $\mu \in M_1(E)$.

Ce théorème permet notamment de traiter le cas de l'équation (3.42) bruitée par une force de la forme (3.43). Cependant les conditions liées (b) et (c) peuvent être modulées selon les modèles : par exemple, quitte à renforcer (c) en $\mathbb{E}|W| < \infty$, on peut affaiblir (b) en

$$\exists 0 < C_0 \leq R_0, C_0 > 0, \quad t.q \ \ |S(x)| \leq |x| - C_0, \quad \forall |x| \geq R_0.$$

En fait, le Théorème 3.7.2 est valide dans le cas plus général d'un espace métrique (E, d) complet et séparable.

3.8 Présentation des résultats du chapitre 6

3.8.1 Généralités sur l'équation de Burgers

En 1948, Burgers se propose d'étudier de manière intensive l'équation parabolique non linéaire suivante comme modèle simple pour la turbulence. On considère un fluide confiné dans le segment $[0, 1]$, l'évolution du champ des vitesses $X(t, \xi)$ est donné par

$$\frac{dX(t, \xi)}{dt} = \Delta X(t, \xi) + \frac{1}{2} D_\xi X^2(t, \xi), \quad t \geq 0,\ \xi \in [0, 1]. \tag{3.47}$$

Cette équation est complétée par des conditions aux bords homogènes de type Dirichlet $X(t,0) = X(t,1) = 0$ et par une condition initiale $X(0,\xi) = x(\xi), \xi \in [0,1]$.

La quantité réelle $X(t,\xi)$ représente la vitesse au temps t du fluide se trouvant à l'endroit ξ. C'est une première approximation pour les équations de Navier-Stokes, puisqu'elle représente de la manière la plus simple l'équilibre entre le processus de convection non linéaire $\frac{1}{2}D_\xi X^2(t)$ et le processus de dissipation $\Delta X(t)$ (On a écrit l'équation avec une viscosité égale à 1).

De nombreux auteurs [5, 6] ont suggéré d'utiliser l'équation de Burgers stochastique comme modèle simple pour la turbulence. Les simulations numériques donnent des résultats intéressants : pour des conditions initiales distinctes on observe une convergence rapide des moyennes en temps vers une valeur commune. Ce chapitre est consacré à l'étude de cette convergence par des méthodes analogues aux chapitres 3 et 4.

Il s'agit en fait d'un travail antérieur, dans lequel certaines étapes de la preuve de notre version du Théorème 3.2.5 sont détaillées, notamment l'obtention de la borne inférieure uniforme. De plus, un résultat de convergence des approximations fini-dimensionnelles est établi. Ces points, ainsi que l'étude du modèle en elle-même, justifient ce chapitre additionnel.

Avant cela, rappelons que pour l'étude déterministe de cette EDP, un outil puissant est la transformée de Hopf-Cole : elle ramène l'équation (3.47) en une équation linéaire du type de la chaleur. Cependant, dans le cas stochastique, cet outil est difficile à utiliser : il transforme le bruit additif en un bruit multiplicatif. On renvoie au travail de Dermoune [16] pour cette approche différente de la nôtre.

3.8.2 Les résultats connus

On considère ici l'équation *stochastique* de Burgers en dimension 1, sur le segment $[0,1]$, qui prend la forme

$$dX(t) = \left(\Delta X(t) + \frac{1}{2}D_\xi X^2(t)\right)dt + Q^{\frac{1}{2}}dW(t); \quad X(0) = x \tag{3.48}$$

avec les conditions de Dirichlet aux bords.

Précisons d'abord les termes qui apparaissent dans cette équation. La condition initiale vérifie $x \in H$ où on note $H := L^2(0,1)$ l'espace des fonctions de carré intégrable équipé de la norme $\|\cdot\|_2$. Le laplacien Δ est un opérateur symétrique non borné défini sur son domaine $D(\Delta)$ dans H, qui est l'intersection de deux espaces de Sobolev (voir [1] pour les définitions)

$$D(\Delta) = \left\{u \in H^2(0,1) : u(0) = u(1) = 0\right\} = H^2(0,1) \cap H_0^1(0,1)$$

On note D_ξ ou ∇ la dérivée par rapport à la variable d'espace. Le processus

$W(t)$ est un Wiener cylindrique sur H qui admet la représentation

$$W(t) = \sum_{k\geq 1} \beta_k(t) e_k(x) \tag{3.49}$$

où les β_k sont des mouvements Browniens réels standards indépendants sur $(\Omega, \mathcal{F}, \mathbb{P})$, et $(e_k)_k$ est la base orthonormée qui diagonalise Δ dans H, dont l'expression explicite est $e_k(x) = \sqrt{2/\pi}\sin(k\pi x)$ pour $k \geq 1$. L'opérateur linéaire symétrique $Q \in L(H)$ vérifie, dans toute la suite, les deux hypothèses suivantes :

(i) Tout d'abord, Q est un opérateur à trace, i.e

$$tr(Q) < \infty \tag{3.50}$$

où $tr(Q) := \sum_{k\geq 1}\langle Qe_k, e_k\rangle_H$, ce qui signifie que l'énergie injectée dans le système est finie. Ainsi $Q^{\frac{1}{2}}W(t)$ prend un sens dans H.

(ii) De plus, Q vérifie la condition de non dégénerescence suivante

$$Im\left((-\Delta)^{-\frac{\delta}{2}}\right) \subset Im\left(Q^{\frac{1}{2}}\right) \text{ pour } \frac{1}{2} < \delta < 1. \tag{3.51}$$

Par exemple, dans le cas simple $Q^{\frac{1}{2}}e_k = \sigma_k e_k$ où $\sigma_k \in \mathbb{R}$, et comme en dimension 1 les valeurs propres de Δ sont les $\lambda_k = k^2\pi^2$, les conditions (3.50) et (3.51) sont vérifiées dès que $c/k < |\sigma_k| \leq C/k^{\frac{1}{2}+\varepsilon}$ pour deux constantes positives c et C et un petit $\varepsilon > 0$. On remarque d'ailleurs, du fait de la dimension, l'exposant 2 qui distingue (3.51) des conditions (3.31) qui jouent le rôle analogue pour Navier-Stokes.

On dénote $S(t)$ le semigroupe du Laplacien sur H, i.e $S(t) = e^{t\Delta}$, et on considère les solutions au sens intégral suivant.

Définition 3.8.1. *On dit que $X \in C([0,T], L^2(0,1))$ est une solution mild de l'équation (3.48) si $X(t)$ est adapté à $(\mathcal{F}_t)$, la tribu de l'information du processus de Wiener jusqu'au temps t et pour tout réel $t \geq 0$, on a*

$$X(t) = S(t)x + \int_0^t S(t-s)\frac{1}{2}D_\xi X^2(s)ds + \int_0^t S(t-s)GdW(s)$$

pour toute condition initiale $x \in H$, $\mathbb{P}$-presque sûrement.

On rappelle qu'en dimension 1, l'espace de Sobolev H_0^1 (voire [1]) est donné par

$$H_0^1 = \{x : [0,1] \to \mathbb{R};\ x \text{ abs. continue}, x(0) = x(1) = 0,\ D_\xi x \in H\}.$$

Le théorème suivant reprend des travaux réalisés autour de l'équation (3.48) durant les quinze dernières années notamment par Da Prato, Debussche, Gatarek, Temam, Goldys et Maslowski (voir [7, 9, 10, 28]).

Théorème 3.8.2. *Sous les conditions (3.50) et (3.51),*

(1) pour tout $T > 0$, l'équation (3.48) admet une unique solution mild, qui vérifie

$$X \in C\left([0,T], L^2(0,1)\right) \cap L^2\left([0,T], H_0^1\right).$$

(2) Le processus solution est de Markov, et son semigroupe de transition P_t sur H admet une mesure invariante.

(3) Enfin, le semigroupe P_t est topologiquement irréductible (TI) et possède la propriété forte de Feller (FF) sur H. En particulier, la mesure invariante est unique. On la note μ par la suite.

Contrairement à l'équation de Navier-Stokes, on a directement les propriétés du semigroupe sur l'espace de base H, ce qui rend l'étude un peu plus directe.

3.8.3 Les résultats obtenus

Comme on l'a déjà souligné, les propriétés (FF) et (TI) impliquent que l'entropie de Donsker Varadhan J vérifie $[J = 0] = \{\mu\}$, l'unique mesure invariante, ainsi que $J(\nu) < +\infty \implies \nu \ll \mu$ pour $\nu \in M_1(H)$.

Elle permettent aussi d'identifier certaines transformées de Legendre avec J (voir Appendix B dans [45]). Ainsi, pour établir dans un premier temps la borne supérieure pour la topologie faible, on peut utiliser une version du théorème de Gärtner-Ellis (voir le Théorème 7.2.1 dans l'Annexe). Il s'agit donc de montrer une propriété de tension exponentielle, à l'aide des estimations a priori suivantes

Théorème 3.8.3. *Pour $0 < \lambda_0 < \frac{\pi^2}{2||Q||}$, le processus N_t à valeurs dans $\mathbb{R}^+$ défini par*

$$N_t := \exp\left(\lambda_0 \int_0^t ||\nabla X(s)||_2^2 ds - \lambda_0 tr(Q)t\right) e^{\lambda_0 ||X(t)||_2^2}$$

est une surmartingale. On a en particulier

$$\mathbb{E}^x e^{\lambda_0 \int_0^t ||\nabla X(s)||_2^2 ds} \leq e^{\lambda_0 tr(Q)t} e^{\lambda_0 ||x||_2^2}, \quad \forall x \in H \tag{3.52}$$

Pour établir ce résultat, on travaille sur les approximations fini-dimensionnelles de la solution données par

$$dX_n(t) = \left(\Delta X_n(t) + \frac{1}{2}\Pi_n D_\xi (X_n)^2(t)\right) dt + G_n dW(t) \ ; \ X_n(0) = \Pi_n x \tag{3.53}$$

où Π_n est la projection othogonale sur $H_n := Vect(e_1, \ldots, e_n)$, l'espace vectoriel engendré par les n premiers vecteurs propres, et $G_n := \Pi_n Q^{\frac{1}{2}}$. N'ayant pas trouvé de résultat de convergence dans la littérature pour le passage à la limite qui donne (3.52), on a prouvé le résultat suivant.

Théorème 3.8.4. *La solution X_n de (3.53) converge vers la solution X de l'équation (3.48) dans l'espace $C([0,T];H)$ et dans l'espace $L^2([0,T];H_0^1)$ presque sûrement.*

Comme pour l'équation de Navier-Stokes, le coté droit de (3.52) incite à introduire, pour $0 < \lambda_0 < \frac{\pi^2}{2\|Q\|}$ et $L > 1$, l'ensemble de mesures initiales

$$\mathcal{M}_{\lambda_0,L} := \left\{ \nu \in M_1(H) \ / \ \int_H \exp(\lambda_0 \|x\|_2^2)\nu(dx) \leq L \right\}$$

sur lequel on établit d'abord la tension exponentielle uniforme et la borne supérieure faible suivante.

Proposition 3.8.1. *Pour tout $\varepsilon > 0$, il existe un compact $K = K_\varepsilon$ de $M_1(H)$ pour la topologie faible tel que*

$$\limsup_{t\to\infty} \frac{1}{t} \log \sup_{\nu \in \mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(L_t \notin K) \leq -\frac{1}{\varepsilon}.$$

Ainsi, pour tout fermé F de $M_1(H)$ muni de la topologie faible $\sigma(M_1(H), C_b(H))$,

$$\limsup_{t\to\infty} \frac{1}{t} \log \sup_{\nu \in \mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(L_t \in F) \leq -\inf_F J.$$

La borne inférieure uniforme est le point délicat. On reprend les notations du Théorème 3.2.5. Soit $\Theta := C(\mathbb{R}^+, H)$ équipé de $\tau_p := \sigma(M_1(\Theta), \cup_{t\geq 0} b\mathcal{F}_t^0)$ où $b\mathcal{F}_t^0$ est l'ensemble des fonctions sur Θ, qui sont mesurables pour $\mathcal{F}_t^0$ et bornées. On rappelle que R_T est la mesure empirique au niveau processus (3.8) et $H : M_1(\Theta) \to [0,\infty]$ l'entropie de niveau 3 définie par (3.9) et on présente maintenant une borne inférieure générale, établie par Wu [45], et rappelée dans le chapitre 7 (Théorème 7.1.1). Pour tout ouvert O dans $(M_1(\Theta), \tau_p)$,

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}_x(R_t \in O) \geq -\inf_O H, \quad \mu - \text{pp } x \in H \tag{3.54}$$

où μ est l'unique mesure invariante. On peut en déduire une borne inférieure uniforme, en se servant d'une borne supérieure. Pour cela on établit en effet la

Proposition 3.8.5. *On suppose que la borne supérieure uniforme suivante est vérifiée : pour tout τ-fermé F dans $M_1(H)$,*

$$\limsup_{T\to\infty} \frac{1}{T} \log \sup_{\nu \in \mathcal{M}} \mathbb{P}_\nu(L_T \in F) \leq -\inf_F J. \tag{3.55}$$

Si J est une bonne fonction de taux sur $(M_1(H), \tau)$, et que la borne générale (3.54) est valide, alors on a la borne inférieure uniforme : pour tout τ-ouvert G dans $M_1(H)$,

$$\liminf_{T\to\infty} \frac{1}{T} \log \inf_{\nu \in \mathcal{M}} \mathbb{P}_\nu(L_T \in G) \geq -\inf_G J.$$

Cependant la Proposition 3.8.1 ne suffit pas, puisqu'il faut une borne supérieure pour la topologie τ. Il faut donc prouver (3.55). Sous les conditions (FF) et (TI), d'après le Théorème B.5 dans [45] et Lemme 7.2.2 (voir annexe), il se trouve que les conditions de retour du Théorème 3.2.5 sont suffisantes.

Il reste à exhiber des compacts de l'espace $H := L^2(0,1)$ en utilisant que l'injection $H_0^1 \subset L^2$ est compacte (c'est un cas particulier du théorème de Rellich, cf [1]). Des candidats naturels sont les boules de l'espace H_0^1, et les estimations de la solution pour la norme $|x|_{H_0^1} = \|\nabla x\|_2$ donnée par le Théorème 3.8.3 permettent de s'en sortir.

Finalement, le PGD de la définition 3.2.4 s'ensuit sur $\mathcal{M} := \mathcal{M}_{\lambda_0,L}$. On montre que J vérifie aussi $J(\nu) < +\infty \implies \nu(H_0^1) = 1$ et $\int_{H_0^1} \|\nabla x\|_2^2 d\nu < +\infty$. Enfin, la topologie τ permet encore des applications à certaines quantités physiques réelles...

Bibliographie

[1] R. A. Adams, *Sobolev spaces*, Academic Press, New-York 1975.

[2] J.M. Bismut, *Large Deviations and the Malliavin Calculus*, Birkhauser (1984).

[3] J. Bricmont, A. Kupiainen, R. Lefevere, Exponential Mixing of the 2D stochastic Navier-Stokes Dynamics , *Com. Math. Phys,* **230***, 87-132 (2002).*

[4] A. Bensoussan, R. Temam, Equations Stochastiques du Type Navier-Stokes , *J. Funct. Anal.,* **13***, 195-222 (1973).*

[5] D.H. Chambers, R.J. Adrian, P. Moin, D.S. Stewart, H.J. Sung, Karhunen-Loeve expansion of Burgers model of turbulence, *Phys. Fluids* **31** *(9), 2573-2582 (1988).*

[6] H. Choi, R. Temam, P. Moin, J. Kim, Feedback control for unsteady flow and its application to Burgers equation *J. Fluid Mech.* **253***, 509-543 (1993).*

[7] G. Da Prato, A. Debussche, Differentiability of the transition semigroup of the Stochastic Burgers equation, and application to the corresponding Hamilton Jacobi equation, *Atti Accad. Naz Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei Mat. Appl.* **9***, 267-277 (1998).*

[8] G. Da Prato and A. Debussche, Ergodicity for the 3D stochastic Navier-Stokes equations, *J. Math. Pures. Appl,* **82***, 877-947 (2003).*

[9] G. Da Prato, A. Debussche and R. Temam, Stochastic Burgers equation, *Nonlinear Analysis* **1***, 389-402 (1994).*

[10] G. Da Prato, D. Gatarek, Stochastic Burgers equation with correlated noise, *Stochastics and Stochastic Reports* **52***, 29-41 (1995).*

[11] C.R. Doering and J.D. Gibbon, *Applied analysis of the Navier-Stokes equations*, Cambridge University Press (1995).

[12] G. Da Prato and J. Zabczyk, *Stochastic equation in infinite dimensions*, Cambridge University Press (1992).

[13] G. Da Prato and J. Zabczyk, *Ergodicity for Infinite Dimensional Systems*, Cambridge University Press (1996).

[14] A. Debussche, C. Odasso, Markov solutions for the 3D stochastic Navier-Stokes equations with state dependant noise *to appear in Journal of Evolution Equations, (2005).*

[15] A. Dembo and O. Zeitouni, *Large Deviations Techniques and applications, 2nd edition*, Springer-Verlag (1998).

[16] A. Dermoune, Around the Stochastic Burgers equation, *Stoch. Analysis and Applic.,* **15(3)***, 295-311 (1997).*

[17] J.D. Deuschel and D.W. Stroock, *Large Deviations*, Pure Appl. Math., Vol. 137, Academic Press, San Diego (1989).

[18] M.D Donsker and S.R.S Varadhan, Asymptotic evalutation of certain Markov process expectations for large time, I-IV, *Comm. Pur. Appli. Math,* **28***, 1-47 and 279-301 (1975) ;* **29***, 389-461 (1976) ;* **36***, 183-212 (1983).*

[19] S.N. Ethier, T.G. Kurz, *Markov Processes Characterization and convergence*, John Wiley and Sons (1986).

[20] K.D. Elworthy, X.M. Li, Formulae for the derivatives of heat semigroups, *J. Funct. Anal.* **125***, 252-286 (1994).*

[21] Weinan E, J.C. Mattingly, Y.G. Sinai, Gibbsian dynamics and ergodicity for the stochastic forced Navier-Stokes equation, *Comm. Math. Phys.,* **224***, (1) (2002).*

[22] B. Ferrario, Ergodic results for stochastic Navier-Stokes Equations, *Stochastics. Stoch. Rep.,* **60***, 271-288 (1997).*

[23] B. Ferrario, Stochastic Navier-Stokes Equations : Analysis of the Noise to Have a Unique Invariant Measure, *Annali di Matematica pura ed applicata, (IV), Vol. CLXXVII, 331-347 (1999).*

[24] F. Flandoli, Dissipativity and invariant measures for stochastic Navier-Stokes Equations, *NoDEA,* **1***, 403-428 (1994).*

[25] F. Flandoli, B. Maslowski, Ergodicity of the 2D Navier-Stokes Equation Under Random Perturbation, *Communications in Mathematical Physics,* **171***, 119-141 (1995).*

[26] F. Flandoli, M. Romito, Markov selections for the 3D stochastic Navier-Stokes equations, *to appear (2006).*

[27] M.I. Freidlin, A.D. Wentzell *Random Perturbations of Dynamical Systems*, Springer-Verlag (1984).

[28] B. Goldys, B. Maslowski, Exponential ergodicity for stochastic Burgers and 2D Navier-Stokes equation, *J. Funct. Anal.,* **226***, Nr 1, 230-255 (2005).*

[29] M. Hairer, J.C. Mattingly, Ergodicity of the 2D Navier-Stokes Equations with Degenerate Stochastic Forcing, *to appear in Ann. Math, (2004).*

[30] M. Hairer, J.C. Mattingly, Spectral Gaps in Wasserstein distances and the 2D Navier-Stokes Equations, *preprint, (2006).*

[31] A.N. Kolmogorov, The local stucture of turbulence in incompressible viscous fluide for very large Reynolds numbers,*C. R. Acad. Sc. URSS* **30***, 301-5.*

[32] S. Kuksin, A. Shirikyan, Stochastic Dissipative PDE's and Gibbs Measures, *Comm. Math. Phys.,* **213***, 291-330, (2000).*

[33] S. Kuksin, A. Shirikyan, A Coupling approach to randomly forced non linear PDE's. I, *Comm. Math. Phys.,* **221***, 351-366 (2001).*

[34] B. Maslowski, On probability distributions of solutions of semilinear stochastic evolution equations, *Stochastics. Stoc. Rep.* **45***, 17-44 (1993).*

[35] N. Masmoudi, L.S. Young, Ergodicity of the 2D-Navier-Stokes equations with Degenerate Stochastic Forcing, *Comm. Math. Phys.,* **227***, 461-481, (2002).*

[36] J.C. Mattingly, *The stochastically-forced Navier Stokes equations : energy estimates and phase space contraction*, P.H.D thesis, Princeton University, (1998).

[37] J.C. Mattingly, E. Pardoux, Malliavin calculus for the stochastic 2D Navier-Stokes equations, *to appear in Comm. Pure Appl. Math (2004).*

[38] G. Métivier, Valeurs propres d'opérateurs définis par la restriction de systèmes variationnels à des sous-espaces, *J. Math. Pures et Appl. 133-156 (1978).*

[39] S.P Meyn and R.L. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag (1993).

[40] C. Odasso, Exponential mixing for the 3D stochastic Navier-Stokes equations, *preprint (2005).*

[41] A. Shirikyan, Exponential mixing for 2D Navier-Stokes Equations Perturbed by an Unbounded Noise, *J. math. fluid mech.,* **6***, 169-193 (2004).*

[42] R. Temam, *Navier-Stokes Equations*, North-Holland (1977).

[43] R. Temam *Navier-Stokes Equations and Nonlinear Functional Analysis, second edition*, S.I.A.M, Philadelphia (1995).

[44] L. Wu, Grandes déviations pour les processus de Markov essentiellement irréductibles, II temps continu, *C.R.Acad.Sci. Sér. I* **314***, 941-946 (1992).*

[45] L. Wu, Uniformly integrable operators and large deviations for Markov processes, *J.Funct.Anal.* **172***, 301-376 (2000).*

[46] L. Wu, Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems, *S.P.A.,* **91***, 205-238 (2001).*

Chapitre 4

A large deviation principle for a 2D stochastic Navier-Stokes equation

Chapitre rédigé en anglais.

Article accepté par Stochastic Processes and their Applications

Abstract : In this paper one specifies the ergodic behavior of the 2D-stochastic Navier-Stokes equation by giving a Large Deviation Principle for the occupation measure for large time. It describes the exact rate of exponential convergence. The considered random force is non-degenerate and compatible with the strong Feller property.

Keywords : Stochastic Navier-Stokes Equation, Large Deviations, Occupation Measure.

AMS Subject Classification 2000: 60F10, 60J35, 35Q30, 76D06.

4.1 Introduction and Results

Let us introduce the two-dimensional incompressible Navier-Stokes equation (NSE in short), which describes the evolution of an incompressible fluid. It is most frequently written in terms of the velocity field u at each point ξ in the domain. Let D be a bounded domain of $\mathbb{R}^2$ with smooth boundary ∂D, we consider the equation

$$\frac{du(t,\xi)}{dt} + (u(t,\xi).\nabla)u(t,\xi) - \Delta u(t,\xi) + \nabla P(t,\xi) = g(\xi) + \eta(t,\xi)\ , \tag{4.1}$$

for $t \geq 0, \xi \in D$, and subject to the incompressibility condition

$$\text{div } u(t,\xi) = 0 \quad , \quad t \geq 0 \quad , \quad \xi \in D \ ,$$

the boundary condition

$$u(t,\xi) = 0 \quad , \quad t \geq 0 \quad , \quad \xi \in \partial D$$

and the initial condition $u(\xi, 0) = x(\xi)$ for $\xi \in D$. In (4.1), the function g is a deterministic external forcing and η a random forcing taking form (4.5) detailed later. For simplicity, we have written the equation in dimensionless form, and with the viscosity equal to 1. It is also possible to work with periodic boundary conditions.

In the usual way, by applying to (4.1) the projection to the linear space of divergence free vector fields (often called Leray projector), the pressure P disappears from the equations. Let $\mathcal{V}$ be the space of $\mathcal{C}^\infty$ 2-dimensional vector fields $u(\xi)$ on D with compact support strictly contained in D, and satisfying div $u(\xi) = 0$. We denote by H (respectively V) its closure in the L^2 topology (respectively in the H^1 topology). According to the classical theory of Navier-Stokes equations, we have

$$\begin{array}{rcl} H & = & \{u \in [L^2(D)]^2 \text{ s.t. div } u = 0, \quad \gamma_\nu(u) = 0\} \\ V & = & \{u \in [H^1_0(D)]^2 \text{ s.t. div } u = 0\} \end{array}$$

where $\gamma_\nu(u)$ is a trace that coincides with $u \cdot \nu$ for smooth u, ν being the outer normal to ∂D (see for example the book of Temam [26, Chap.1]).

Moreover $|\cdot|$ and $\langle \cdot, \cdot \rangle$ stand for the norm and inner product in H. Identifying H with its dual H' and identifying H' with a subspace of V' (the dual space of V) we have $V \subset H \subset V'$, and we also denote by $\langle \cdot, \cdot \rangle$ the duality between V and V'. Let us define the linear operator A in H by the formula

$$Au = -P_{div}\Delta u \quad , \quad \forall u \in D(A) = \left(H^2(D)\right)^2 \cap V$$

and the bilinear operator $B : V \times V \to V'$ by $B(u,v) = P_{div}(u \cdot \nabla)v$ where P_{div} is the L^2 projection operator onto the space H of divergence-free vector field. The space V coincides with $D\left(A^{\frac{1}{2}}\right)$ and is endowed with the norm $|x|_V = |A^{\frac{1}{2}}x|$. The unbounded linear operator A is closed, positive and selfadjoint in H, with compact inverse A^{-1}. Following classical spectral theory, we denote by $0 < \lambda_1 \leq \lambda_2 \leq ...$ the eigenvalues of A and by $e_1, e_2, ..$ a corresponding complete orthonormal system of eigenvectors. Finally we can define the fractional powers A^α and their domains, the spaces $D\left(A^\alpha\right)$ equipped with the norm $|u|_\alpha := |A^\alpha u|$, that correspond to Sobolev spaces $\left[H^{2\alpha}(D)\right]^2$ with the suitable conditions. We remark in particular that $D\left(A^\alpha\right)$ is dense in $D\left(A^\beta\right)$ for $\alpha > \beta \geq 0$, and that for any $\alpha > 0$,

$$|x| \leq \frac{1}{{\lambda_1}^\alpha}|A^\alpha x|. \tag{4.2}$$

The incompressibility condition implies for any u, v, z in V

$$\langle B(u,v), v\rangle = 0 \quad , \quad \langle B(u,v), z\rangle = -\langle B(u,z), v\rangle. \tag{4.3}$$

By applying to each term of the NSE the projection operator P_{div}, we formally rewrite the system (4.1) in the abstract form :

$$dX(t) + AX(t)dt + B\left(X(t), X(t)\right)dt = fdt + P_{div}\eta(t,\xi)dt \; ; \; X(0) = x \tag{4.4}$$

where $X(t)$ is identified with $u(t,\cdot)$ and $f = P_{div}g$ (the irrotational components of f and η are absorbed in the term $\nabla P(t,\xi)$, see [26]).

In mathematical literature, it is common to assume the random force $P_{div}\eta(t,\xi)dt$ to be random fields that are smooth in x, while as a function of time t they are white noises (see [7] for example). Since we are interested in the long time behavior of equation (4.4), both the forcing terms are assumed to be stationary in order to have an autonomous system (i.e $f \in H$ do not depend on the time variable t, whereas the white noise is by definition a stationary process).

Let us describe the form of the noise. We assume that

$$P_{div}\eta(t,\xi)dt = GdW(t) \tag{4.5}$$

where $W(t)$ is a standard cylindrical Wiener process in H (see [7]) defined on a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $G : H \to H$ is a bounded linear operator satisfying

$$D\left(A^{2\alpha}\right) \subset Im(G) \subset D\left(A^{\frac{1}{2}+\varepsilon}\right) \quad , \quad \text{for some } \frac{1}{4} < \alpha < \frac{1}{2} \quad , \quad \varepsilon \in \left(0, 2\alpha - \frac{1}{2}\right]. \tag{4.6}$$

Here, $Im(G)$ is the range of the operator G. Roughly speaking, the first embedding in (4.6) means that the noise is not too degenerate, and the second implies that $tr(G^*G) < \infty$ (i.e, the energy injected by the random force is finite) and also gives us more spatial regularity for the solution to (4.4).

The stochastic NSE has been intensively studied since the work of Bensoussan and Temam [2]. Here we adopt the approach of generalized solutions given by Flandoli [13] (see for instance Flandoli and Gatarek [14] for solutions in law called martingale solutions). On a fixed probability space he built an associated Markovian semigroup of transition with an invariant measure. Under a condition of non-degenerance of type (4.6) for the noise, the uniqueness of the invariant measure was first showed by Flandoli and Maslowski [15] and in a more classical way by Ferrario [12]. Goldys and Maslowski [17] established recently the exponentially fast convergence of transition measures to the invariant measure. More references on the degenerated noise case will be presented in Remarks 4.1.6.

Under (4.6), it is known (see next section for more details and references) that the solution $X(t)$ of (4.4) is a Markov process with a unique invariant measure μ

supported by $D(A^\alpha)$. By the uniqueness (see [7]), μ is ergodic in the sense that

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \Psi\left(X(t,x)\right)dt = \int \Psi d\mu \quad \mathbb{P}\text{-a.s.}$$

for all initial conditions x and all continuous and bounded functions Ψ.

In the sequel, $\mathbb{P}_x$ is the law on $C(\mathbb{R}^+, H)$ of the Markov process with $x \in H$ as initial state, and for any initial measure ν on H, let $\mathbb{P}_\nu(\cdot) := \int_H \mathbb{P}_x(\cdot)\nu(dx)$. Our aim is to establish the large deviation principle (LDP in short) for the occupation measure L_t of the solution X to (4.4) given by

$$L_t(A) := \frac{1}{t}\int_0^t \delta_{X(s)}(A)ds, \quad \forall A \in \mathcal{B}(H)$$

where δ_a is the Dirac measure at a, and $\mathcal{B}(H)$ the Borelian σ-field in H. Notice that L_t is a random measure on H but in fact supported on $D(A^\alpha)$, because of the regularity of X, given by Theorem 4.2.2 below. The LDP for empirical measure is one of the strongest ergodic result for the long time behavior of Markov processes. This is a traditional subject in probability since the pioneering work of Donsker and Varadhan [10], however in our infinite dimensional setting their assumptions are not satisfied (see [8, 9] for an introduction to Large Deviation theory).

Let us begin with some necessary definitions. For $E = H$ or $E = D(A^\alpha)$, let $M_1(E)$ be the space of probability measures (resp. $M_b(E)$ the space of signed σ-additive measures of bounded variation) on E equipped with the Borel σ-field $\mathcal{B}$. On the space $M_b(E)$ (or $M_1(E)$ its subspace), we consider $\sigma(M_b(E), b\mathcal{B}(E))$, the so called τ-topology of convergence against measurable and bounded functions which is much stronger than the usual weak convergence topology $\sigma(M_b(E), C_b(E))$. The duality relation between $\nu \in M_b(E)$ and $\Psi \in b\mathcal{B}(E)$ will be denoted by

$$\nu(\Psi) := \int_E \Psi d\nu.$$

It is time to state our main result.

Theorem 4.1.1. *Let $f \in H$ and $\frac{1}{4} < \alpha < \frac{1}{2}$ a fixed number such that (4.6) holds (throughout this paper). Let $0 < \lambda_0 < \frac{\lambda_1}{2\|Q\|}$, where $\|Q\|$ is the norm of $Q := GG^*$ as an operator in H and*

$$\mathcal{M}_{\lambda_0,L} := \left\{\nu \in M_1(H) \ / \ \int_H e^{\lambda_0|x|^2}\nu(dx) \le L\right\}. \tag{4.7}$$

The family $\mathbb{P}_\nu(L_T \in \cdot)$ as $T \to +\infty$ satisfies the large deviation principle (LDP) with respect to the topology τ, with speed T and the rate function J, uniformly for any initial measure ν in $\mathcal{M}_{\lambda_0,L}$ where $L > 1$ is any fixed number. Here the rate function $J : M_1(H) \to [0, +\infty]$ is the level-2 entropy of Donsker-Varadhan defined by (4.14) below. More precisely we have:

i) J is a good rate function on $M_1(H)$ equipped with the topology τ of the convergence against bounded and borelian functions, i.e., $[J \le a]$ is τ-compact for every $a \in \mathbb{R}^+$.

ii) for all open set G in $M_1(H)$ with respect to the topology τ,

$$\liminf_{T\to\infty} \frac{1}{T} \log \inf_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(L_T \in G) \ge -\inf_G J$$

iii) for all closed set F in $M_1(H)$ with respect to the topology τ,

$$\limsup_{T\to\infty} \frac{1}{T} \log \sup_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(L_T \in F) \le -\inf_F J \ .$$

Furthermore, we have for μ the invariant measure, and $\forall \nu \in M_1(H)$,

$$J(\nu) < +\infty \implies \nu \ll \mu,\ \nu(V) = 1 \text{ and } \int_V |A^{\frac{1}{2}}x|^2 d\nu < +\infty. \tag{4.8}$$

The LDP w.r.t. the topology τ is much stronger than that w.r.t. the usual weak convergence topology as in Donsker-Varadhan [10]. Indeed, this theorem and the estimate (4.23) below have interesting consequences for which the topology τ is crucial. For instance, LDP can be deduced for many non-continuous physical observables of the system such as $|x|_V := |A^{\frac{1}{2}}x|$, the Sobolev norm (which is not continuous on H).

Corollary 4.1.2. *Let $(\mathbb{B}\ ,\ \|\cdot\|_{\mathbb{B}})$ a separable Banach space, and $f : D\left(A^{\frac{1}{2}}\right) \to \mathbb{B}$ a measurable function bounded on the balls $\{x \text{ s.t } |A^{\frac{1}{2}}x| \le R\}$ for any $R > 0$, and satisfying*

$$\lim_{|A^{\frac{1}{2}}x|\to\infty} \frac{\|f(x)\|_{\mathbb{B}}}{|A^{\frac{1}{2}}x|^2} = 0. \tag{4.9}$$

Then $\mathbb{P}_\nu(L_T(f) \in \cdot)$ satisfies the LDP on $\mathbb{B}$ with speed T and the rate function I_f given by

$$I_f(y) = \inf\{J(\nu);\quad J(\nu) < +\infty,\ \nu(f) = y\},\ \forall y \in \mathbb{B}$$

uniformly over initial distributions ν in $\mathcal{M}_{\lambda_0,L}$ (for any $L > 1$).

As a particular case of Corollary 4.1.2, we can state the

Proposition 4.1.3. *As $T \to \infty$, the family*

$$\mathbb{P}_\nu\left(\frac{1}{T}\int_0^T X_t dt \in \cdot\right)$$

satisfies a LDP on $D\left(A^{\frac{1}{2}}\right)$ *with speed* T *and the rate function* I *defined by*

$$I(z) = \inf\left\{J(\nu);\quad J(\nu) < +\infty,\ \nu\,(x) = z\right\},\ \forall z \in D\left(A^{\frac{1}{2}}\right)$$

uniformly over initial distributions ν *in* $\mathcal{M}_{\lambda_0,L}$ *(for any* $L > 1$*).*

Remarks 4.1.4. We give now two examples of noises in our class (4.6). Let us first represent the cylindrical Wiener process $W(t)$ as a series with respect to the system $(e_k)_k$ which diagonalizes A on its domain, and define $Ge_k = \sigma_k e_k$, so that

$$GW(t) = \sum_{k=1}^{\infty} \sigma_k \beta_k(t) e_k$$

where $(\beta_k)_{k\in\mathbb{N}}$ is a family of independent real valued standard Brownian motions. The condition (4.6) is

$$\frac{c}{k^{2\alpha}} \leq \sigma_k \leq \frac{C}{k^{\frac{1}{2}+\varepsilon}}$$

for two positive constants c and C, and k large enough since $\lambda_k \sim k$ as $k \to \infty$. Hence, the cylindrical Wiener process with values in $D\left(A^{2\alpha}\right)$, that is when $\sigma_i \lambda_i^{2\alpha} = 1$, is allowed.

A more general example of noise for which our assumption holds for $\frac{1}{4} < \alpha < \frac{1}{2}$ fixed is $G := A^{-\beta}L$ where L is any linear bounded and invertible operator on H and $\frac{1}{2} < \beta \leq 2\alpha$.

Remarks 4.1.5. The class (4.7) of initial distributions for the uniform LDP is sufficiently rich. For example, choosing L large enough, it includes all the Dirac probability measures δ_x with x in any ball of H.

Remarks 4.1.6. During the last years, a lot of progresses have been done in the treatment of very degenerated noises (acting only on a finite number of modes as in Kolmogorov's turbulence theory). In this case Bismut-Elworthy-Li formula become irrelevant, and the Strong Feller property does not hold. However a careful analysis of the dynamic allow to obtain uniqueness of invariant measure. We refer to the works of Weinan E, Mattingly and Sinai (see [11, 23]) where only a finite number N of low modes is forced where N depends on the viscosity. More recently, Hairer, Mattingly and Pardoux (see [20]) have removed this last dependence of N on the viscosity and have established the uniqueness when only four lowest modes are forced. Degenerated kick noises have also been considered by Kuksin and Shirikyan among others (see [22]). For all thoses cases, we believe that the LDP w.r.t the topology τ is false. The LDP w.r.t the weak convergence topology in that degenerated noise case is an interesting open problem. Finally the 3D case is much more delicate. But it seems possible, by selecting solutions to build a Makovian semigroup with strong

Feller and irreducibility properties (see Da-Prato, Debussche [6] and Odasso [24]). It is hoped that our method will make it possible to treat this $3D$-case with a non degenerate noise.

Remarks 4.1.7. In another direction, a Wentzell-Freidlin type large deviation principle was proved by Chang [4] for the paths of the solution when the magnitude of the additive noise tends to zero. This result is extended to the multiplicative noise case by Sritharan and Sundar [25] (see also the recent works of Collina, Livi, Mazzino [5] and Amirdjanova, Xiong [1]).

This paper is organized as follows. In section 2, we recall results on existence and uniqueness of solution, and invariant probability measure for the equation (4.4). In section 3, we present some general facts about Large Deviations for strongly Feller and topologically irreducible Markov processes. Then, in section 4, we prove a useful exponential estimate for the solution, and we do some comments on the rate function which governs the LDP. We establish first this LDP on $D(A^\alpha)$ in section 5, and we extend it on H in section 6. Finally, Proposition 4.1.3 and Corollary 4.1.2 are investigated in section 7.

4.2 Existence and uniqueness results for the solution and the invariant measure

Following the literature ([12, 13, 15, 17] among many others), we say that a progressively measurable process $X(t)$ is a generalized solution of equation (4.4) if

$$X \in C([0,T],H) \cap L^2\left([0,T], D(A^{\frac{1}{4}})\right) \qquad \mathbb{P}-a.s.$$

and the equation is satisfied $\mathbb{P}$-a.s. in the weak sense

$$\begin{aligned} & \langle X(t), y\rangle + \int_0^t \langle X(s), Ay\rangle ds - \int_0^t \langle B\left(X(s), y\right), X(s)\rangle ds \\ = \; & \langle x, y\rangle + t\langle f, y\rangle + \langle GW(t), y\rangle \end{aligned}$$

for all $t \in [0,T]$, $y \in D(A)$ and the initial condition $x \in H$. This definition is justified by the properties (4.3) of the non linearity B, and the Sobolev continuous embedding $D(A^{\frac{1}{4}}) \subset [L^4(D)]^2$ since

$$\langle B\left(X(s), y\right), X(s)\rangle \le C|A^{\frac{1}{2}}y|\; |X(s)|^2_{[L^4(D)]^2} \le C|A^{\frac{1}{2}}y|\; |A^{\frac{1}{4}}X(s)|^2.$$

Hence all the terms make sense.

Flandoli [13] proved for the first existence of a solution under the weaker assumption $Im(G) \subset D(A^{\frac{1}{4}+\varepsilon})$: the classical definition of solution was not adapted here because of the low regularity of the noise. However, under our condition (4.6), the noise is more regular and his result can be read as

Theorem 4.2.1. [13] *Assume that (4.6) holds for $\frac{1}{4} < \alpha < \frac{1}{2}$.*

For all $x \in H,\ f \in D\left(A^{-\frac{1}{2}}\right)$, there exists a unique generalized solution X^x of the equation (4.4) such that $\mathbb{P}$-a.e.

$$X^x \in C([0,T],H) \cap L^2\left([0,T], D(A^{\frac{1}{2}})\right). \tag{4.10}$$

and $X^x - Z \in L^2\left([0,T], D(A^{\frac{1}{2}})\right)$ where Z is the solution to the auxiliary Ornstein-Uhlenbeck equation

$$dZ(t) + AZ(t) = GdW(t).$$

The family of solutions X^x for $x \in H$ forms a Markov family which admits an invariant measure μ.

The following step consists in analyzing whether μ is unique. For this purpose topological irreducibility and strong Feller property were investigated. We recall first the definitions.

Denote by E a generic space. Given the solution X^x, a E-valued continuous process starting from $x \in E$, the transition functions $P(t,x,\Gamma) := \mathbb{P}\left(X^x(t) \in \Gamma\right)$ are well defined for any $t \leq T, x \in E$ and Γ any measurable subset of E. The topological irreducibility in E means that $P(t,x,O) > 0$ for some $t > 0, x \in E$ and any non-empty open subset O of E, and P_t is strongly Feller if $P_t : b\mathcal{B}(E) \to C_b(E)$.

In the case of the stochastic Navier-Stokes equation, Flandoli and Maslowski [15] proved the topological irreducibility in H and the Strong Feller property in $D(A^{\frac{1}{4}})$. They obtained thus the uniqueness. But, for the investigation of a large deviation principle, we have a powerful criterion if the semigroup is topologically irreducible and strongly Feller on the same space. So, our beginning is the following theorem for solutions starting from a regular initial condition due to Ferrario [12].

Theorem 4.2.2. *[12] Assume that (4.6) holds for $\frac{1}{4} < \alpha < \frac{1}{2}$.*

(i) If $x \in D\left(A^\alpha\right),\ f \in D\left(A^{\alpha-\frac{1}{2}}\right)$, the unique solution X^x of equation (4.4) given by Theorem 4.2.1 satisfies in fact $\mathbb{P}$-a.e.

$$X^x \in C\left([0,T], D\left(A^\alpha\right)\right) \cap L^2\left([0,T], D(A^{\frac{1}{2}})\right) \cap L^{\frac{4}{1-2\alpha}}\left([0,T], D\left(A^{\frac{1}{4}+\frac{\alpha}{2}}\right)\right)$$

and also $X^x - Z \in L^2\left([0,T], D\left(A^{\alpha+\frac{1}{2}}\right)\right)$ $\mathbb{P}$-*a.e.*

(ii) The process (X^x) is Markovian, and its transition probability $P_t f(x) := \mathbb{E} f(X^x_t)$ is topologically irreducible and strongly Feller in $D\left(A^\alpha\right)$. In particular, the invariant measure μ is unique.

(iii) Moreover, for every $t_0 > 0$, and every $x \in H$, the corresponding solution satisfies $\mathbb{P}$-a.e. $X^x \in C([t_0, T], D(A^\alpha))$.

In fact the original assumption of Ferrario was

$$D\left(A^{2\alpha}\right) \subset Im(G) \subset D\left(A^{\frac{1}{4}+\frac{\alpha}{2}+\varepsilon'}\right) \tag{4.11}$$

for $\alpha \in [\frac{1}{4}, \frac{1}{2})$ and some $\varepsilon' > 0$. However, the second embedding in (4.11) is clearly implied by the second embedding in (4.6). But our condition implies also that the energy injected in the system is finite. More precisely, we recall the

Lemma 4.2.3. *If the linear and continuous operator $G : H \to H$ satisfies for some $\varepsilon > 0$*

$$Im(G) \subset D(A^{\frac{1}{2}+\varepsilon}) \tag{4.12}$$

then the symmetric nonnegative operator $Q := GG^*$ is of trace class.

The above finite trace property is crucial in the application of Itô's formula for establishing our exponential estimates, and implies the usual regularity (4.10) for the solution.

4.3 General results about large deviations

In this section, we recall general results on the Large Deviation Principle for strong Feller and topologically irreducible Markov processes. We follow [27] and [28].

4.3.1 Notations and entropy of Donsker-Varadhan

Here we consider a general E-valued continuous Markov process,

$$\left(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, (X_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E}\right)$$

whose semigroup of Markov transition kernels is denoted by $(P_t(x, dy))_{t\geq 0}$, where : $\Omega = C(\mathbb{R}^+, E)$ is the space of continuous functions from $\mathbb{R}^+$ to E equipped with the compact convergence topology; the natural filtration is $\mathcal{F}_t = \sigma(X_s, 0 \leq s \leq t)$ for any $t \geq 0$ and $\mathcal{F} = \sigma(X_s, 0 \leq s)$. As usual, the law of the Markov process with initial state x in E is $\mathbb{P}_x$, and for any initial measure ν on E, let $\mathbb{P}_\nu(\cdot) = \int_E \mathbb{P}_x(\cdot)\nu(dx)$.

The empirical measure of level-3 (or process level) is given by

$$R_t := \frac{1}{t}\int_0^t \delta_{\theta_s X} ds$$

where $(\theta_s X)_t = X_{s+t}$ for all $t, s \geq 0$ are the shifts on Ω. Thus, R_t is a random element of $M_1(\Omega)$, the space of all probability measure on Ω.

The level-3 entropy functional of Donsker-Varadhan $H : M_1(\Omega) \to [0, +\infty]$ is defined by

$$H(Q) := \begin{cases} \mathbb{E}^{\bar{Q}} h_{\mathcal{F}_1}\left(\bar{Q}_{\omega(-\infty,0]}; \mathbb{P}_{w(0)}\right) & \text{if } Q \in M_1^s(\Omega) \\ +\infty & \text{otherwise} \end{cases} \tag{4.13}$$

where: $M_1^s(\Omega)$ is the space of those elements in $M_1(\Omega)$ which are moreover stationary; $\bar{Q}$ is the unique stationary extension of $Q \in M_1^s(\Omega)$ to $\bar{\Omega} := C(\mathbb{R}, E)$; the filtration is extended on $\bar{\Omega}$ with $\mathcal{F}_t^s = \sigma(X(u); s \leq u \leq t)$, $\forall s, t \in \mathbb{R}$; finally $\bar{Q}_{X(-\infty,t]}$ is the regular conditional distribution of $\bar{Q}$ knowing $\mathcal{F}_t^{-\infty}$ and $h_{\mathcal{G}}(\nu, \mu)$ is the usual relative entropy or Kullback information of ν with respect to μ restricted on the σ-field $\mathcal{G}$, given by

$$h_{\mathcal{G}}(\nu, \mu) := \begin{cases} \int \frac{d\nu}{d\mu}|_{\mathcal{G}} \log\left(\frac{d\nu}{d\mu}|_{\mathcal{G}}\right) d\mu, \text{ if } \nu \ll \mu \text{ on } \mathcal{G} \\ +\infty \text{ otherwise.} \end{cases}$$

Now, the level-2 entropy functional $J : M_1(E) \to [0, \infty]$ which governs the LDP in our main result is

$$J(\beta) = \inf\{H(Q) \mid Q \in M_1^s(\Omega) \text{ and } Q_0 = \beta\}, \ \forall \beta \in M_1(E), \tag{4.14}$$

where $Q_0(\cdot) = Q(X(0) \in \cdot)$ is the marginal law at $t = 0$.

Proposition 4.3.1. *For our model, $J(\nu) < +\infty \implies \nu \ll \mu$. Furthermore, a necessary and sufficient condition for $J(\nu) = 0$ is $\nu = \mu$.*

Proof. Here we take $E := D(A^\alpha)$, where X_t is strongly Feller and topologically irreducible by Theorem 4.2.2. Consider ν such that $J(\nu) < \infty$. By definition, there exists some $Q \in M_1^s(\Omega)$ such that $Q_0 = \nu$, and $H(Q) < \infty$ (see the expression (4.13) giving the Level-3 entropy).

For such Q and all $t > 0$, we have by stationarity (see [27, App. B])

$$\begin{aligned} H(Q) &= \mathbb{E}^{\bar{Q}} h_{\mathcal{F}_1}\left(\bar{Q}_{X(-\infty,0]}; \mathbb{P}_{X_0}\right) \\ &= \frac{1}{t}\mathbb{E}^{\bar{Q}} h_{\mathcal{F}_t}\left(\bar{Q}_{X(-\infty,0]}; \mathbb{P}_{X_0}\right). \end{aligned}$$

By Jensen inequality we obtain

$$\frac{1}{t}\mathbb{E}^{\bar{Q}} h_{\mathcal{F}_t}\left(\bar{Q}_{X(-\infty,0]}; \mathbb{P}_{X_0}\right) \geq \frac{1}{t} h_{\mathcal{F}_t}\left(Q; \mathbb{P}_\nu\right).$$

Then, noting that entropy of marginal measures is not larger than global entropy,

$$\begin{aligned} H(Q) &\geq \frac{1}{t} h_{\sigma(X_t)}\left(Q; \mathbb{P}_\nu\right) \\ &\geq \frac{1}{t} h_{\mathcal{B}(E)}\left(\nu; \nu P_t\right) \end{aligned}$$

and taking infinimum over such Q, we get

$$J(\nu) \geq \frac{1}{t} h_{\mathcal{B}(E)}(\nu; \nu P_t). \tag{4.15}$$

So the Kullback information of ν with respect to νP_t is finite, which implies by definition that $\nu \ll \nu P_t$. Since P_t is topologically irreducible and strongly Feller, all the measures $P_t(x, dy), t > 0, x \in E$ are equivalent to μ (see [7, p. 42]), and we have

$$\nu P_t = \int_E P_t(x, .)\nu(dx) \ll \mu.$$

Thus $\nu \ll \nu P_t \ll \mu$.

If the probability measure ν is such that $J(\nu) = 0$ then $h_{\mathcal{B}(E)}(\nu; \nu P_t) = 0$ using (4.15). By the well known property of the Kullback information, we obtain $\nu = \nu P_t$ for every $t \geq 0$. Finally, the uniqueness of the invariant measure for P_t in Theorem 4.2.2 implies $\nu = \mu$ and the proof is finished. □

4.3.2 The hyper-exponential recurrence criterion

In that case we have the following criterion of the so-called hyper-exponential recurrence established by Wu [28, Theorem 2.1] (for a general polish space E).

Theorem 4.3.2. *[28] Let $\mathcal{A} \subset M_1(E)$ and assume that*

$$P_t \textit{ is strong Feller and topologically irreducible on } E. \tag{4.16}$$

If $\forall \lambda > 0$ there exists some compact $K \subset\subset E$, such that

$$\sup_{\nu \in \mathcal{A}} \mathbb{E}^\nu e^{\lambda \tau_K} < \infty \tag{4.17}$$

and

$$\sup_{x \in K} \mathbb{E}^x e^{\lambda \tau_K^{(1)}} < \infty \tag{4.18}$$

where $\tau_K := \inf\{t \geq 0 \ s.t \ X_t \in K\}$ and $\tau_K^{(1)} := \inf\{t \geq 1 \ s.t \ X_t \in K\}$, then the family $\mathbb{P}_\nu(L_t \in \cdot)$ satisfies the LDP on $M_1(E)$ w.r.t to the τ-topology with the rate function J defined by (4.14), and uniformly for initial measures ν in the subset $\mathcal{A}$.

More precisely, the three properties hold:

(a1) $J : M_1(E) \to [0, +\infty]$ is inf-compact w.r.t the τ-topology

(a2)(the lower bound) for any τ-open G in $M_1(E)$,

$$\liminf_{T \to \infty} \frac{1}{T} \log \inf_{\nu \in \mathcal{A}} \mathbb{P}_\nu(L_T \in G) \geq -\inf_G J \tag{4.19}$$

(a3)(the upper bound) for any τ-closed F in $M_1(E)$,

$$\limsup_{T\to\infty} \frac{1}{T} \log \sup_{\nu\in\mathcal{A}} \mathbb{P}_\nu(L_T \in F) \le -\inf_F J. \tag{4.20}$$

This theorem is in fact a slight extension of the result in [28] to a uniform LDP over a non-empty family of initial measures. Let us recall briefly the main steps of the proof and the corresponding references (see however [19] for a complete proof). At first, a pointwise level-3 lower bound can be deduced from the properties (4.16) via the notion of μ-essential irreducibility (see [27]). This pointwise lower bound yields the uniform lower bound (4.19) if the uniform upper bound (4.20) is satisfied (as in [18]).

So, essential part of the proof is the uniform upper bound (4.20) for the strong topology τ. Indeed, the upper bound for the weak topology may be proved by the exponential tightness implied by (4.17) and (4.18) (see section 4.4.1), but the strong Feller property is crucial for the τ topology. By an extension of the Gartner-Ellis theorem (see [27]), it is sufficient to prove that $\forall\ (f_n) \in B_b(E)$ decreasing to zero pointwisely on E, we have

$$\lim_{n\to\infty} \limsup_{T\to\infty} \frac{1}{T} \log \sup_{\nu\in\mathcal{A}} \mathbb{E}^\nu \int_0^T f_n(X_s)ds = 0.$$

This last assertion follows from the Markovian and the strong Feller property, also from (4.17) and (4.18) and can be proved as in [28]. Actually this last point is a problem for establishing the LDP for degenerated noise with a unique invariant measure.

4.4 Exponential estimates for the solution and some comments on the rate function J

In this section we establish the following crucial exponential estimates for the solution.

Proposition 4.4.1. *For any fixed $0 < \lambda_0 < \frac{\lambda_1}{2\|Q\|}$ and any $x \in H$, the process X satisfies*

$$\mathbb{E}^x \exp\left(\lambda_0|X(t)|^2 + \int_0^t \lambda_0 |A^{\frac{1}{2}}X(s)|^2 ds\right) \le e^{\left(\lambda_0\left(Tr(Q)+\frac{|f|^2}{\lambda_1-2\|Q\|\lambda_0}\right)t\right)} e^{\lambda_0|x|^2}.$$

In particular, the following estimates hold

$$\mathbb{E}^x \exp\left(\lambda_0|X(t)|^2 ds\right) \le \exp\left(\lambda_0 t\left(tr(Q) + \frac{|f|^2}{\lambda_1 - 2\|Q\|\lambda_0}\right)\right) e^{\lambda_0|x|^2} \tag{4.21}$$

and

$$\mathbb{E}^x \exp\left(\lambda_0 \int_0^t |A^{\frac{1}{2}}X(s)|^2 ds\right) \le \exp\left(\lambda_0 t\left(tr(Q) + \frac{|f|^2}{\lambda_1 - 2\|Q\|\lambda_0}\right)\right) e^{\lambda_0|x|^2}. \quad (4.22)$$

Moreover, for any fixed $L > 1$, we have

$$\sup_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{E}^\nu \exp\left(\lambda_0 \int_0^t |A^{\frac{1}{2}}X(s)|^2 ds\right) \le e^{\left(\lambda_0 t\left(tr(Q)+\frac{|f|^2}{\lambda_1-2\|Q\|\lambda_0}\right)\right)} L \quad (4.23)$$

where $\mathcal{M}_{\lambda_0,L}$ is the set of initial measures defined by (4.7).

Before proving this property at the end of this section, we first give some consequences of these estimates, and some comments about the entropy J of Donsker-Varadhan.

4.4.1 First consequences of the exponential estimates

The first one is the

Corollary 4.4.2. *Under the estimate (4.22), the family of law $\mathbb{P}_\nu(L_t \in \cdot)$ is uniformly exponentially tight over $\mathcal{M}_{\lambda_0,L}$. More precisely, for any $\varepsilon > 0$, there is some compact subset $K = K_\varepsilon$ in $M_1(H)$ in the weak convergence topology such that*

$$\limsup_{t\to\infty} \frac{1}{t} \log \sup_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(L_t \notin K) \le -\frac{1}{\varepsilon}.$$

Consequently for any closed set F in $M_1(H)$ equipped with the weak convergence topology $\sigma\left(M_1(H), C_b(H)\right)$, we have

$$\limsup_{t\to\infty} \frac{1}{t} \log \sup_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(L_t \in F) \le -\inf_F J \quad (4.24)$$

where the entropy of Donsker-Varadhan $J : \nu \in M_1(H) \to J(\nu) \in [0, +\infty]$ satisfies

$$\lambda_0 \int_H |A^{\frac{1}{2}}x|^2 d\nu \le J(\nu) + \lambda_0\left(tr(Q) + \frac{|f|^2}{\lambda_1 - 2\|Q\|\lambda_0}\right) \quad (4.25)$$

for any $0 < \lambda_0 < \frac{\lambda_1}{2\|Q\|}$.

Proof. The proof of the exponential tightness from (4.22) is given in [18, Prop. 5.1]. According to the general theory, it yields the upper bound (4.24) by using a general weak upper bound on compact subsets (see [8]).

For the proof of (4.25), let us recall the definition of a Cramer functional

$$\Lambda^0(V) := \sup_{x\in H} \limsup_{t\to\infty} \frac{1}{t} \log \mathbb{E}^x \exp\left(tL_t(V)\right), \quad \forall V \in b\mathcal{B}(H) \tag{4.26}$$

and the definition of its Legendre transformation:

$$(\Lambda^0)^*(\nu) = \sup_{V\in b\mathcal{B}(H)} \int_H V d\nu - \Lambda^0(V), \quad \forall \nu \in M_1(H)). \tag{4.27}$$

It is known that $(\Lambda^0)^* = J$ (see for instance [27, prop. B.13]). Let us consider the function $V_n(x) := \lambda_0 |A^{\frac{1}{2}} x|^2 \wedge n$ which is bounded and measurable on H (Here $a \wedge b$ is the minimum of two real numbers a and b). For $0 < \lambda_0 < \frac{\lambda_1}{2\|Q\|}$, by the definitions (4.26), (4.27), and the exponential estimate (4.22), we obtain

$$\begin{aligned} \nu(V_n) &\le (\Lambda^0)^*(\nu) + \Lambda^0(V_n) \\ &\le J(\nu) + \lambda_0 \left(tr(Q) + \frac{|f|^2}{\lambda_1 - 2\|Q\|\lambda_0} \right). \end{aligned}$$

Hence, we obtain (4.25) by Fatou's lemma.

□

In fact, this kind of estimates provides us an alternative way for proving the existence of an invariant measure based on the large deviations theory. More precisely, we have the

Corollary 4.4.3. *Assume that a Feller-Markov process X on H satisfies the exponential estimate (4.22), then X admits at least one invariant measure.*

Proof. It is sufficient to prove that for any $a > 0$, the subset $[J \le a]$ is tight. In that case, by Prokorov's criterion, its closure is compact in $M_1(H)$ w.r.t the weak topology. Hence, the l.s.c function $J : M_1(H) \to [0, +\infty]$ admits compact level subsets, w.r.t the weak topology $\sigma\left(M_1(H), C_b(H)\right)$. Moreover, considering the closed subset $F = M_1(H)$ in the good upper bound (4.24), we obtain the existence of $\nu \in M_1(H)$ satisfying $J(\nu) = 0$, so ν is an invariant measure (as in the proof of Proposition 4.3.1).

Now, for any $a > 0$, let us show that the tightness of $[J \le a]$ is a simple consequence of (4.25). Let $\varepsilon > 0$ fixed, and consider the finite number

$$M_{a,\varepsilon} := \frac{a + \lambda_0 \left(tr(Q) + \frac{|f|^2}{\lambda_1 - 2\|Q\|\lambda_0} \right)}{\lambda_0 \varepsilon}.$$

By the compact embedding $D(A^{\frac{1}{2}}) \subset H$, the subset

$$K_\varepsilon = \left\{ \; x \in D(A^{\frac{1}{2}}) \; s.t. \; |A^{\frac{1}{2}} x|^2 \le M_{a,\varepsilon} \; \right\}$$

is compact in H, and by using (4.25) we obtain for all β in $[J \leq a]$,

$$\begin{aligned}\beta(K_\varepsilon^c) &\leq \int_{K_\varepsilon^c} \frac{|A^{\frac{1}{2}}x|^2}{M_{a,\varepsilon}}\, d\beta(x) \leq \frac{1}{M_{a,\varepsilon}} \int_H |A^{\frac{1}{2}}x|^2\, d\beta(x) \\ &\leq \frac{1}{M_{a,\varepsilon}} \frac{a + \lambda_0 \left(tr(Q) + \frac{|f|^2}{\lambda_1 - 2\|Q\|\lambda_0} \right)}{\lambda_0} \\ &\leq \varepsilon\end{aligned}$$

and so $[J \leq a]$ is tight. □

In the following paragraph we focus on the entropy J which governs the LDP.

4.4.2 Some comments on the entropy J defined by (4.14)

In fact $J(\nu)$ admits a closed form only in the case when the unique invariant measure μ is known, and the Markov process X_t is symmetric w.r.t μ. For our model, Theorem 4.1.1 describes the exact rate of exponential convergence, but the expression of this exact rate given by J is more qualitative than quantitative. How to obtain estimates on $J(\nu)$ is an important and very interesting question. Usually, we can proceed by using functional inequalities such as logarithmic Sobolev or spectral gap inequalities as in Deuschel-Stroock [9]. Unfortunately, for the 2D Stochastic Navier-Stokes equations, those inequalities are actually unknown (see however the very recent works of Goldys-Maslowski [17] and Hairer-Mattingly [21] for the existence of a spectral gap in a space of weighted bounded or weighted Lipschitz functions, which is different from the Poincaré inequality). In this section, we consider the case $E := H$.

At first, under the Feller assumption, we know that (see Lemma B.7 in [27])

$$J(\nu) = \sup \left\{ - \int \frac{\mathcal{L}u}{u} d\nu \; ; \; 1 \leq u \in D_e(\mathcal{L}) \right\}, \qquad \nu \in M_1(H) \tag{4.28}$$

where $D_e(\mathcal{L})$ is the extended domain of the generator $\mathcal{L}$ of P_t in $C_b(H)$. We recall that $u \in D_e(\mathcal{L})$ and $v := \mathcal{L}u$ if $u \in C_b(E)$ and there exists $v \in C_b(H)$ such that $P_t u - u = \int_0^t P_s v ds$, for all $t \geq 0$. For the 2D-stochastic Navier-Stokes equation, we recall also that $\mathcal{L}$ is given by

$$\mathcal{L}f(x) := \frac{1}{2} tr(GG^* D^2 f)(x) \; + \; \langle -Ax - B(x,x) + f, \nabla^H f(x) \rangle \tag{4.29}$$

at least for f cylindrical, i.e $f(x) = g\left(\langle x, e_1\rangle, \langle x, e_2\rangle, \dots, \langle x, e_n\rangle\right)$. In this expression, we denote by ∇^H the gradient in H, and $D^2 f := \left(\partial_{e_i}\partial_{e_j} f\right)_{i,j \geq 1}$. Since f is cylindrical, the gradient $\nabla^H f(x)$ is in $H^k := D(A^{\frac{k}{2}})$, for any $k \geq 0$ and the left-hand side in (4.29) is well defined by $\langle B(x,x), \nabla^H f(x)\rangle = -\langle B(x, \nabla^H f(x)), x\rangle$ and the inequality

$$|\langle B(x, \nabla^H f(x)), x\rangle| \leq C|x|^2 |A\nabla^H f(x)| \left(1 + \log \frac{|A^{\frac{3}{2}}\nabla^H f(x)|^2}{\lambda_1 |A\nabla^H f(x)|^2}\right)^{\frac{1}{2}}$$

established for instance in [16, p 100].

In this paragraph, we introduce some class of measures $\mu^h \in M_1(H)$ for which it is possible to give a more explicit form than (4.14) or (4.28) for $J(\mu^h)$. Here, we assume that $G = G^* = Q^{\frac{1}{2}}$. Let (X_t^x) be the solution of the 2D stochastic Navier-Stokes equation (4.4) with initial position x, defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ and let us consider the Girsanov perturbation defined by : for any $T > 0$, and any $x \in H$,

$$\frac{d\mathbb{Q}_x^h}{d\mathbb{P}}|_{\mathcal{F}_T} = \exp\left(\int_0^T \sqrt{GG^*}\nabla^H h(X_s^x) dW_s - \frac{1}{2}\int_0^T \left|\sqrt{GG^*}\nabla^H h(X_s^x)\right|^2 ds\right). \tag{4.30}$$

In the above expression, we take $h \in C^1(H)$ satisfying $\langle GG^*\nabla^H h, \nabla^H h\rangle \leq C < \infty$ so that

$$\mathbb{E}\exp\left(\frac{1}{2}\int_0^T \left|\sqrt{GG^*}\nabla^H h(X_s^x)\right|^2 ds\right) < \infty, \quad \forall T \geq 0,\ \forall x \in H. \tag{4.31}$$

A simple sufficient condition on h is $h(x) = g\left(\langle x, e_1\rangle, \langle x, e_2\rangle, \ldots, \langle x, e_n\rangle\right)$ where the function $g \in C^1(\mathbb{R}^n)$ has a bounded gradient.

In the case when (4.31) is true, $L_t^x := \int_0^t \sqrt{GG^*}\nabla^H h(X_s^x) dW_s$ is a continuous martingale under $\mathbb{P}$, and also $M_t^x := \exp(L_t^x - \frac{1}{2}\langle L^x\rangle_t)$, the exponential local martingale given in (4.30), becomes a true martingale by Novikov's criterion.

Hence, $(\mathbb{Q}_x^h)_{x\in H}$ given in (4.30) defines a new Markov family with the transition semigroup $Q_t^h f(x) = \mathbb{E}^{\mathbb{Q}_x^h} f(X_t^x)$. By Girsanov's formula, the generator of Q_t^h takes the form $\mathcal{L}_h u = \mathcal{L}u + 2\Gamma(h, u)$ where

$$\Gamma(h, u) = 1/2\ \langle GG^*\nabla^H h, \nabla^H u\rangle$$

is the *carré du champ* of $\mathcal{L}$, and under $\mathbb{Q}_x^h$ the process (X_t^x) satisfies in a weak sense (i.e. in law) the following perturbation of the 2D-stochastic Navier-Stokes equation

$$dX_t + AX_t dt + B\left(X_t, X_t\right) dt = f dt + \sqrt{GG^*}\nabla^H h\left(X_t\right) dt + G d\tilde{W}_t \quad , \quad X_0 = x \tag{4.32}$$

where $\tilde{W}_t$ is a cylindrical Wiener process under $\mathbb{Q}_x^h$.

For the existence and the uniqueness of an invariant measure $\mu^h \in M_1(H)$ for Q_t^h, let us first give the

Lemma 4.4.4. *Let $0<\delta<\frac{\lambda_1}{4\|Q\|}$ and $h\in C^1(H)$ such that*

$$\langle GG^*\nabla^H h(x),\nabla^H h(x)\rangle \le C<\infty, \tag{4.33}$$

then there is some constant $K(\delta)>0$ such that

$$\mathbb{E}^{\mathbb{Q}^h_x}\exp\left(\delta\int_0^t |A^{\frac12}X^x_s|^2ds\right)\le e^{tK(\delta)}e^{2\delta|x|^2}. \tag{4.34}$$

Proof. By using (4.30) and Cauchy-Schwartz inequality, we obtain

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^h_x}\exp\left(\delta\int_0^t |A^{\frac12}X^x_s|^2ds\right) &= \mathbb{E}^{\mathbb{P}}\exp\left(\delta\int_0^t |A^{\frac12}X^x_s|^2ds\right)\exp\left(L^x_t-\frac12\langle L^x\rangle_t\right)\\
&= \mathbb{E}^{\mathbb{P}}\exp\left(\delta\int_0^t |A^{\frac12}X^x_s|^2ds+\frac12\langle L^x\rangle_t\right)\exp\left(L^x_t-\langle L^x\rangle_t\right)\\
&\le \left(\mathbb{E}^{\mathbb{P}}\exp\left(2\delta\int_0^t |A^{\frac12}X^x_s|^2ds+\langle L^x\rangle_t\right)\right)^{\frac12}
\end{aligned}$$

since the exponential local martingale $\exp\left(2L^x_t-2\langle L^x\rangle_t\right)$ is a supermartingale. Hence, noting that $\langle L\rangle_t\le Ct$ by (4.33), and using estimate (4.22) with $\lambda_0=2\delta$, we obtain

$$\mathbb{E}^{\mathbb{Q}^h_x}\exp\left(\delta\int_0^t |A^{\frac12}X^x_s|^2ds\right) \le \exp\left(\delta t\left(tr(Q)+\frac{|f|^2}{\lambda_1-4\|Q\|\delta}\right)\right)e^{\frac{Ct}{2}}e^{\delta|x|^2}$$

i.e the estimate (4.34). □

For $x,y\in H$, the following control is well known (see [26])

$$|X^x_t-X^y_t|\le|x-y|\exp\left(C\int_0^t |A^{\frac12}X^x_s|^2ds\right)$$

which implies the convergence in probability $X^y_t\to X^x_t$ and $f(X^y_t)\to f(X^x_t)$ for any function $f\in C_b(H)$, as $y\to x$. In particular, the Feller property for P_t follows by Lebesgue Theorem. We prove now that Q^h_t is also a Feller semigroup.

We must show that for any sequence $x_n\to x$ in H, and for any $f\in C_b(H)$, the convergence $Q^h_tf(x_n)\to Q^h_tf(x)$ holds. By using the density given in (4.30), it is equivalent to prove that $\mathbb{E}^{\mathbb{P}}f(X^{x_n}_t)M^{x_n}_t\to\mathbb{E}^{\mathbb{P}}f(X^x_t)M^x_t$. The quadratic variation process of the $\mathbb{P}$-martingale $L^{x_n}_t-L^x_t$ satisfies

$$\begin{aligned}
\langle L^{x_n}-L^x\rangle_t &= \int_0^t\langle GG^*(\nabla^Hh(X^{x_n}_s)-\nabla^Hh(X^x_s)),\nabla^Hh(X^{x_n}_s)-\nabla^Hh(X^x_s)\rangle ds\\
&\le 4Ct
\end{aligned}$$

by polarization under our assumption (4.33). Hence, the convergence $X^{x_n}_t\to X^x_t$ in probability implies the convergence in L^2 for the martingale $L^{x_n}_t\to L^x_t$, and in particular the convergence in probability for the exponential martingale $M_t^{x_n} \to M_t^x$. Since $\mathbb{E}^{\mathbb{P}} M_t^{x_n} = \mathbb{E}^{\mathbb{P}} M_t^x = 1$, and $M_t^{x_n}, M_t^x \geq 0$ we obtain the convergence in L^1 for $M_t^{x_n} \to M_t^x$ by a well known lemma. Finally, since $f(X_t^{x_n}) \to f(X_t^x)$ in probability and is bounded, we obtain $\mathbb{E}^{\mathbb{P}} f(X_t^{x_n}) M_t^{x_n} \to \mathbb{E}^{\mathbb{P}} f(X_t^x) M_t^x$ as desired.

Hence the exponential estimate (4.34) implies the existence of an invariant measure for the Feller semigroup Q_t^h by Corollary 4.4.3. Moreover, by (4.30), we know that $Q_t^h(x,\cdot) \sim P_t(x,\cdot) \sim P_t(y,\cdot) \sim Q_t^h(y,\cdot)$. So, the semigroup Q_t^h is regular and its invariant measure μ^h is unique (see Doob's theorem in [7]).

In that case, we have the following simple expression for $J(\mu^h)$, where μ^h can be seen as the unique invariant measure for the solution to the equation (4.32).

Proposition 4.4.5. *For $h \in C^1(H)$ such that $\langle GG^* \nabla^H h(x), \nabla^H h(x)\rangle \leq C < \infty$, we have*

$$J(\mu^h) \; = \; \frac{1}{2}\int_H \langle GG^* \nabla^H h, \nabla^H h\rangle d\mu^h \; = \; \int_H \Gamma(h,h) d\mu^h$$

Proof. For the "$\geq$", let us consider a nice approximating sequence of cylindrical functions h_n for h such that $h_n(x) \to h(x)$, $\nabla^H h_n(x) \to \nabla^H h(x)$ and $|\nabla^H h_n(x)| \leq M$. Since $\mathcal{L}e^{h_n} = e^{h_n}\left(\mathcal{L}h_n + \Gamma(h_n,h_n)\right)$, we have using (4.28)

$$\begin{aligned} J(\mu^h) \; &\geq \; -\int \frac{\mathcal{L}e^{h_n}}{e^{h_n}} d\mu^h \\ &= \; -\int \mathcal{L}_h h_n d\mu^h + \int 2\Gamma(h,h_n) - \Gamma(h_n,h_n) d\mu^h \\ &= \; \int \left[2\Gamma(h,h_n) - \Gamma(h_n,h_n)\right] d\mu^h \end{aligned}$$

because μ^h is the invariant measure for the semigroup Q_t^h generated by $\mathcal{L}_h$. Now, letting $n \to \infty$, we obtain $J(\mu^h) \geq \int \Gamma(h,h) d\mu^h$ by dominated convergence.

For the "$\leq$", we denote by $\mathbb{Q}^h_{\mu^h}$ the law of the unique stationary Markov process with μ^h as initial distribution and the transition semigroup Q_t^h. By our assumption on h, we have

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}^h_{\mu^h}} \langle L\rangle_t \; &= \; 2\mathbb{E}^{\mathbb{Q}^h_{\mu^h}} \int_0^t \Gamma(h,h)(X_s) ds \\ &= \; 2t \int \Gamma(h,h) d\mu^h \\ &\leq \; Ct < \infty \end{aligned}$$

where the second equality follows from the stationarity of $\mathbb{Q}^h_{\mu^h}$. Hence, $L_t^x - \langle L^x\rangle_t$ being a $\mathbb{Q}^h_{\mu^h}$-local martingale by Girsanov, is in fact a true $\mathbb{Q}^h_{\mu^h}$-martingale, and the definition (4.13) of the level-3 entropy gives

$$H(\mathbb{Q}^h_{\mu^h}) \; = \; \mathbb{E}^{\mathbb{Q}^h_{\mu^h}} \log M_1$$

$$
\begin{aligned}
&= \mathbb{E}^{\mathbb{Q}^h_{\mu^h}}\left(L_1^x - \frac{1}{2}\langle L^x\rangle_1\right) \\
&= \frac{1}{2}\mathbb{E}^{\mathbb{Q}^h_{\mu^h}}\langle L^x\rangle_1.
\end{aligned}
$$

So, again with the stationarity of $\mathbb{Q}^h_{\mu^h}$, we obtain

$$
\begin{aligned}
H(\mathbb{Q}^h_{\mu^h}) &= \frac{1}{2}\mathbb{E}^{\mathbb{Q}^h_{\mu^h}}\int_0^1 \left|\sqrt{GG^*}\nabla^H h(X_s^x)\right|^2 ds \\
&= \frac{1}{2}\int_H \left|\sqrt{GG^*}\nabla^H h(x)\right|^2 d\mu^h
\end{aligned}
$$

Finally, by the definition (4.14) of the level-2 entropy, we have

$$
J(\mu^h) \le H(\mathbb{Q}^h_{\mu^h}) = \frac{1}{2}\int_H \left|\sqrt{GG^*}\nabla^H h(x)\right|^2 d\mu^h = \int \Gamma(h,h)d\mu^h.
$$

□

4.4.3 Proof of Proposition 4.4.1

We finish this section by giving the proof of the exponential estimates. Let us introduce the finite dimensional approximations system associated with equation (4.1). Let Π_n the orthogonal projections on the finite dimensional space spanned by the first n eigenvectors $(e_1, \ldots, e_n)$, and set, for $n \ge 1$,

$$
B_n(x) = \Pi_n B(\Pi_n x, \Pi_n x), \quad G_n = \Pi_n G \Pi_n, \quad f_n = \Pi_n f
$$

and $Q_n = G_n G_n^*$. We will consider the finite dimensional equations

$$
dX_n(t) + AX_n(t)dt + B_n\left(X_n(t)\right) = f_n dt + G_n dW(t)\ ;\ X_n(0) = x_n := \Pi_n x. \tag{4.35}
$$

Note that equation (4.35) is a finite-dimensional stochastic equation. Hence, there exists a solution, and $X_n(t)$ is smooth in space. Moreover, the following convergence was proved by Capinski and Gatarek [3] (see also Goldys and Maslowski [17]).

Theorem 4.4.6. *[17] For any $\delta > 0$, solution X_n of (4.35) converge in distribution to the solution X of (4.4) on the space $C([0,T], H^{-\delta})$, where $H^{-\delta}$ is the dual space of H^δ.*

The first aim of this paragraph is to prove some estimates on X_n, the finite dimensional approximations. Recall that $\langle\ ,\ \rangle$ denote the scalar product in H. Let us apply Itô's formula to the finite dimensional diffusion X_n. Since by (4.3),

$$
\langle B_n\left(X_n(t)\right), X_n(t)\rangle = 0
$$

we obtain:

$$
\begin{aligned}
d|X_n(t)|^2 &= 2\langle X_n(t), dX_n(t)\rangle + tr(Q_n)dt \\
&= \left[-2|A^{\frac{1}{2}}X_n(t)|^2 + 2\langle X_n(t), f_n\rangle + tr(Q_n)\right] dt + 2\langle X_n(t), G_n dW(t)\rangle.
\end{aligned}
$$

Hence, for $U_n(t) := |X_n(t)| + \int_0^t |A^{\frac{1}{2}}X_n(s)|ds$, it yields

$$
dU_n(t) = \left[-|A^{\frac{1}{2}}X_n(t)|^2 + 2\langle X_n(t), f_n\rangle + tr(Q_n)\right] dt + 2\langle X_n(t), G_n dW(t)\rangle.
$$

Same manner, denoting by $d[U_n, U_n]_t$ the quadratic variation process of U_n, we can also compute with the Itô formula

$$
\begin{aligned}
de^{\lambda_0 U_n(t)} &= e^{\lambda_0 U_n(t)} \left[\lambda_0 dU_n(t) + \frac{\lambda_0^2}{2} d[U_n, U_n]_t\right] \\
&= \lambda_0 e^{\lambda_0 U_n(t)} \left[-|A^{\frac{1}{2}}X_n(t)|^2 + 2\langle X_n(t), f_n\rangle + tr(Q_n) + 2\lambda_0 |G_n^* X_n(t)|^2\right] dt \\
&\quad +2\lambda_0 e^{\lambda_0 |X_n(t)|^2} \langle X_n(t), G_n dW(t)\rangle. \qquad (4.36)
\end{aligned}
$$

The following inequalities are clear

$$
tr(Q_n) \leq tr(Q) \quad , \quad |x_n| \leq |x| \quad , \quad |f_n| \leq |f|. \qquad (4.37)
$$

Moreover, it is easy to see that

$$
|G_n^* X_n(t)|^2 \leq \|Q\| \ |X_n(t)|^2, \qquad (4.38)
$$

and, by Young's inequality, that

$$
2\langle X_n(t), f_n\rangle \leq \varepsilon |X_n(t)|^2 + \frac{|f_n|^2}{\varepsilon}, \quad \forall \varepsilon > 0. \qquad (4.39)
$$

For $\varepsilon > 0$ fixed later, let us estimate the drift of the process

$$
V_n(t) := e^{-\lambda_0\left(Tr(Q) + \frac{|f|^2}{\varepsilon}\right)t} e^{\lambda_0 U_n(t)}.
$$

By Itô Formula and using (4.36), (4.37), (4.38) and (4.39) we have,

$$
\begin{aligned}
dV_n(t) &= e^{-\lambda_0\left(Tr(Q) + \frac{|f|^2}{\varepsilon}\right)t} de^{\lambda_0 U_n(t)} - \lambda_0 \left(Tr(Q) + \frac{|f|^2}{\varepsilon}\right) e^{-\lambda_0\left(Tr(Q) + \frac{|f|^2}{\varepsilon}\right)t} e^{\lambda_0 U_n(t)} dt \\
&\leq \lambda_0 e^{-\lambda_0\left(Tr(Q) + \frac{|f|^2}{\varepsilon}\right)t} e^{\lambda_0 U_n(t)} \left(-|A^{\frac{1}{2}}X_n(t)|^2 + \varepsilon |X_n(t)|^2 + 2\lambda_0 \|Q\| \ |X_n(t)|^2\right) dt \\
&\quad +2\lambda_0 e^{-\lambda_0\left(Tr(Q) + \frac{|f|^2}{\varepsilon}\right)t} e^{\lambda_0 |X_n(t)|^2} \langle X_n(t), G_n dW(t)\rangle.
\end{aligned}
$$

Remarking that, by (4.2), for a constant λ_1 depending on the domain D,

$$|X_n(t)|^2 \leq \frac{|A^{\frac{1}{2}} X_n(t)|^2}{\lambda_1},$$

we obtain

$$\begin{aligned} dV_n(t) &\leq e^{-\lambda_0\left(Tr(Q)+\frac{|f|^2}{\varepsilon}\right)t} e^{\lambda_0 U_n(t)} \left(-|A^{\frac{1}{2}}X_n(t)|^2 \left(1-\frac{\varepsilon+2\lambda_0\|Q\|}{\lambda_1}\right)\right)dt \\ &\quad +2\lambda_0 e^{-\lambda_0\left(Tr(Q)+\frac{|f|^2}{\varepsilon}\right)t} e^{\lambda_0|X_n(t)|^2}\langle X_n(t), G_n dW(t)\rangle. \end{aligned}$$

Hence, for $0<\lambda_0\leq \frac{\lambda_1-\varepsilon}{2\|Q\|}$, the drift of $V_n(t)$ is non positive. More precisely, for all $0<\lambda_0<\frac{\lambda_1}{2\|Q\|}$, the positive number $\varepsilon:=\lambda_1-2\|Q\|\lambda_0$ satisfies the above condition, and it is our choice in (4.39). Thus, we have

$$dV_n(t) \leq 2\lambda_0 \exp\left(-\lambda_0\left(Tr(Q)+\frac{|f|^2}{\varepsilon}\right)t\right) e^{\lambda_0|X_n(t)|^2}\langle X_n(t), G_n dW(t)\rangle.$$

Since $V_n(t)\geq 0$, we obtain by Fatou's lemma $\mathbb{E}^x V_n(t)\leq \mathbb{E}^x V_n(0)$, and this proves in particular the following crucial exponential estimate.

Lemma 4.4.7. *For $0<\lambda_0<\frac{\lambda_1}{2\|Q\|}$ and any x in H, we have*

$$\begin{aligned} &\mathbb{E}^x \exp\left(\lambda_0\int_0^t |A^{\frac{1}{2}}X_n(s)|^2 ds+\lambda_0|X_n(t)|^2\right) \\ &\leq \exp\left(\lambda_0 t\left(tr(Q)+\frac{|f|^2}{\lambda_1-2\|Q\|\lambda_0}\right)\right)e^{\lambda_0|x|^2}. \end{aligned} \tag{4.40}$$

Let us now finish the proof of Proposition 4.4.1 by using Theorem 4.4.6. Since the function

$$F(X):=e^{\left(-\lambda_0 t\left(tr(Q)+\frac{|f|^2}{\lambda_1-2\|Q\|\lambda_0}\right)\right)}\exp\left(\lambda_0\int_0^t|A^{\frac{1}{2}}X(s)|^2ds+\lambda_0|X(t)|^2\right)$$

is lower semi continuous on $C([0,T],H^{-\delta})$ as an increasing limit of the continuous functions

$$F_m(X) := e^{\left(-\lambda_0 t\left(tr(Q)+\frac{|f|^2}{\lambda_1-2\|Q\|\lambda_0}\right)\right)}\exp\left(\lambda_0\int_0^t|A^{\frac{1}{2}}\Pi_m X(s)|^2ds+\lambda_0|\Pi_m X(t)|^2\right),$$

we obtain, using Theorem 4.4.6 for $n\to\infty$ in (4.40),

$$\mathbb{E}^x F(X)\leq \liminf_{n\to\infty}\mathbb{E}^x F(X_n)\leq e^{\lambda_0|x|^2}$$

and the desired estimates (4.21) and (4.22) follow.

$\square$

4.5 The Large Deviation Principle on $M_1(D(A^\alpha))$

The proof of our Theorem 4.1.1 consists in two steps. As a first step, we prove this section the LD principle for initial measures in $E := M_1(D(A^\alpha))$. We finish the proof of Theorem 4.1.1 in the following section by extending the LDP for initial conditions, open and closed subsets in $M_1(H)$, and by establishing the claim (4.8).

The aim of this section is to prove the

Lemma 4.5.1. *Let $f \in H$ and $\frac{1}{4} < \alpha < \frac{1}{2}$ a fixed number such that (4.6) holds. Let $0 < \lambda_0 < \frac{\lambda_1}{2\|Q\|}$, where $\|Q\|$ is the norm of Q as an operator in H and*

$$\Phi(x) = e^{\lambda_0|x|^2},\ \mathcal{M}^*_{\lambda_0,L} := \Big\{\nu \in M_1(D(A^\alpha))\ /\ \int \Phi(x)\nu(dx) \le L\Big\}, \tag{4.41}$$

then the family $\mathbb{P}_\nu(L_T \in \cdot)$ as $T \to +\infty$ satisfies the LDP on $M_1(D(A^\alpha))$ w.r.t the topology τ, with speed T and the rate function J, uniformly for any initial measure in $\mathcal{M}^*_{\lambda_0,L}$ where $L > 1$ is any fixed number, and J is the level-2 entropy of Donsker-Varadhan. More precisely, the statements (i),(ii) and (iii) of Theorem 4.1.1 hold with $M_1(H)$ replaced by $M_1(D(A^\alpha))$.

Proof. By Theorem 4.3.2, since X_t is strongly Feller and topologically irreducible in $D(A^\alpha)$ (Theorem 4.2.2), it is sufficient to establish the estimates (4.17) and (4.18) for our model.

For K, we take

$$K := \Big\{x \in D(A^{\frac{1}{2}})\ s.t\ |A^{\frac{1}{2}}x| \le M\Big\} \tag{4.42}$$

where the real M will be fixed later. Since the embedding $D\left(A^{\frac{1}{2}}\right) \subset D(A^\alpha)$ is compact for $\alpha < \frac{1}{2}$, it is clear that K is a compact subset in $D(A^\alpha)$.

The definition of the occupation measure implies that

$$\mathbb{P}_\nu\left(\tau_K^{(1)} > n\right) \le \mathbb{P}_\nu\left(L_n(K) \le \frac{1}{n}\right) = \mathbb{P}_\nu\left(L_n(K^c) \ge 1 - \frac{1}{n}\right).$$

With our choice for K, we have $L_n(K^c) \le \frac{1}{M^2} L_n(|A^{\frac{1}{2}}x|^2)$. Hence, for any fixed λ_0 such that $0 < \lambda_0 < \frac{\lambda_1}{2\|Q\|}$, we obtain by Chebychev's inequality

$$\begin{aligned}
\mathbb{P}_\nu\left(\tau_K^{(1)} > n\right) &\le \mathbb{P}_\nu\left(L_n(|A^{\frac{1}{2}}x|^2) \ge M^2\left(1 - \frac{1}{n}\right)\right) \\
&\le \exp\left(-n\lambda_0 M^2\left(1 - \frac{1}{n}\right)\right)\mathbb{E}^\nu \exp\left(\lambda_0 \int_0^n |A^{\frac{1}{2}}X(s)|^2 ds\right).
\end{aligned}$$

For any initial measure $\nu \in M_1(D(A^\alpha))$, integrating (4.22) w.r.t $\nu(dx)$, and plugging it in the above estimate yields

$$\mathbb{P}_\nu\left(\tau_K^{(1)} > n\right) \le \nu\left(e^{\lambda_0|\cdot|^2}\right) e^{-n\lambda_0 C}, \quad \forall n \ge 2$$

where

$$C := \frac{M^2}{2} - tr(Q) - \frac{|f|^2}{\lambda_1 - 2\|Q\|\lambda_0}. \tag{4.43}$$

Let $\lambda > 0$ be fixed. By the integration by parts formula, we have

$$\begin{aligned}
\mathbb{E}^\nu e^{\lambda\tau_K^{(1)}} &= 1 + \int_0^{+\infty} \lambda e^{\lambda t}\mathbb{P}_\nu\left(\tau_K^{(1)} > t\right) dt \\
&\le e^{2\lambda} + \sum_{n\ge2} \lambda e^{\lambda(n+1)}\mathbb{P}_\nu\left(\tau_K^{(1)} > n\right) \\
&\le e^{2\lambda}\left(1 + \lambda\ \nu\left(e^{\lambda_0|\cdot|^2}\right)\ \sum_{n\ge2} e^{-n(\lambda_0 C-\lambda)}\right).
\end{aligned}$$

Now, by (4.43), we can choose M such that $\lambda_0 C - \lambda \ge 1$ in the definition (4.42) of K. Then, taking the supremum over $\{\nu = \delta_x, x \in K\}$, we get

$$\sup_{x\in K} \mathbb{E}^x e^{\lambda\tau_K^{(1)}} \le e^{2\lambda}\left(1 + \lambda e^{\frac{\lambda_0 M^2}{\lambda_1}} \sum_{n\ge1} e^{-n(\lambda_0 C-\lambda)}\right) < \infty$$

since for $x \in K$, $|x|^2 \le \frac{|A^{\frac{1}{2}}x|^2}{\lambda_1} \le \frac{M^2}{\lambda_1}$. So the bound (4.17) holds true. We obtain (4.18) in the same way: since $\tau_K \le \tau_K^{(1)}$, we have

$$\begin{aligned}
\sup_{\nu\in\mathcal{M}^*_{\lambda_0,L}} \mathbb{E}^\nu e^{\lambda\tau_K} &\le \sup_{\nu\in\mathcal{M}^*_{\lambda_0,L}} \mathbb{E}^\nu e^{\lambda\tau_K^{(1)}} \\
&\le e^{2\lambda}\left(1 + \lambda L \sum_{n\ge2} e^{-n(\lambda_0 C-\lambda)}\right) \\
&< \infty
\end{aligned}$$

which finishes the proof of the Lemma.

□

4.6 Proof of the Theorem 4.1.1

In this section we finish the proof of Theorem 4.1.1. It remains to extend Lemma 4.5.1 for initial conditions, open and closed subsets in $M_1(H)$, and to establish the claim (4.8). In fact, the first claim "$J(\nu) < +\infty \implies \nu \ll \mu$" in (4.8) was established in Proposition 4.3.1 and the second claim that for ν such that $J(\nu) < \infty$, we have also $\nu(|A^{\frac{1}{2}}x|^2) < \infty$ follows from (4.25).

For the extension of the LDP to $M_1(H)$, a first remark is the

Lemma 4.6.1. *Let $L > 1$, $0 < \lambda_0 < \frac{\lambda_1}{2\|Q\|}$ be fixed numbers and $\nu \in \mathcal{M}_{\lambda_0,L}$, i.e $\nu \in M_1(H)$ and $\int_H e^{\lambda_0|x|^2}\nu(dx) \le L$. Then the probability measure $\tilde{\nu} := \mathbb{P}_\nu(X(1) \in dy)$ satisfies*

(a) $\tilde{\nu}(D(A^\alpha)) = 1$

(b)$\int_H e^{\lambda_0|x|^2}\tilde{\nu}(dx) \le e^{\lambda_0 C}L$, where $C = tr(Q) + \frac{|f|^2}{\lambda_1 - 2\|Q\|\lambda_0}$

In particular $\tilde{\nu} \in \mathcal{M}^*_{\lambda_0, e^{\lambda_0 C}L}$, and $\mathbb{P}_\nu(L_T \circ \theta_1 \in \cdot) = \mathbb{P}_{\tilde{\nu}}(L_T \in \cdot)$.

Proof. Notice that $\tilde{\nu}(dy) = \int_H \mathbb{P}_x(X(1) \in dy)\,\nu(dx)$, and $\mathbb{P}_x(X(1) \in D(A^\alpha)) = 1$, by Theorem 4.2.2, (iii) the statement (a) is clear. For (b), by using the estimate (4.21) in Proposition 4.4.1, we get

$$\begin{aligned}
\int_H e^{\lambda_0|y|^2}\tilde{\nu}(dy) &= \int_H\int_H e^{\lambda_0|y|^2}P_1(x,dy)\nu(dx) = \int_H \mathbb{E}^x e^{\lambda_0|X(1)|^2}\nu(dx)\\
&\le \int_H e^{\lambda_0\left(tr(Q)+\frac{|f|^2}{\lambda_1-2\|Q\|\lambda_0}\right)}e^{\lambda_0|x|^2}\nu(dx)\\
&\le e^{\lambda_0\left(tr(Q)+\frac{|f|^2}{\lambda_1-2\|Q\|\lambda_0}\right)}L.
\end{aligned}$$

□

4.6.1 Extension of the lower bound

Let G be an open subset in $(M_1(H), \tau)$. For any fixed $\beta_0 \in M_1(H)$, we can take a τ neighborhood of β_0 in $M_1(H)$ of form

$$N(\beta_0, \delta) := \{\beta \in M_1(H), |\beta(f_i) - \beta_0(f_i)| < \delta, \forall i = 1...d\}$$

contained in G, where $\delta > 0$, $1 \le d \in \mathbb{N}$ and $f_i \in b\mathcal{B}(H)$. For establishing the lower bound (ii) in Theorem 4.1.1, it is sufficient to establish that for every $\beta_0 \in G$ such that $J(\beta_0) < \infty$,

$$\liminf_{T\to\infty}\frac{1}{T}\log\inf_{\nu\in\mathcal{M}_{\lambda_0,L}}\mathbb{P}_\nu(L_T \in N(\beta_0,\delta)) \ge -J(\beta_0).$$

Notice that for $\nu \in \mathcal{M}_{\lambda_0,L}$,

$$\mathbb{P}_\nu(L_T \in N(\beta_0,\delta)) \ge \mathbb{P}_\nu\left(L_T\circ\theta_1 \in N(\beta_0,\delta/2);\ |L_T\circ\theta_1(f_i) - L_T(f_i)| \le \delta/2,\ \forall i = 1..d\right)$$

and

$$|L_T\circ\theta_1(f_i) - L_T(f_i)| \le \frac{2\|f_i\|_\infty}{T}$$

so, we obtain for $T \geq \frac{4}{\delta} \max_{1\leq i\leq d} \|f_i\|_\infty$

$$\begin{aligned}\mathbb{P}_\nu \left(L_T \in N(\beta_0,\delta)\right) &\geq \mathbb{P}_\nu \left(L_T \circ \theta_1 \in N(\beta_0,\delta/2)\right)\\ &\geq \mathbb{P}_{\tilde{\nu}} \left(L_T \in N(\beta_0,\delta/2)\right).\end{aligned}$$

We conclude by using the uniform lower bound on $\mathcal{M}^*_{\lambda_0, e^{\lambda_0 C}L}$, obtained in the preceding section.

4.6.2 Extension of the upper bound

Let F closed in $(M_1(H),\tau)$ such that $\inf_F J = a > 0$ (else the upper bound is clear). We define

$$F_\delta := \left\{\beta \in M_1(H) \ : \ d_{\|\cdot\|_{var}}(\beta, F) < \delta\right\}$$

where $d_{\|\cdot\|_{var}}(\beta,F) := \inf_{\lambda\in F} \|\beta-\lambda\|_{var}$, and the total variation norm of λ is

$$\|\lambda\|_{var} := \sup_{f\in b\mathcal{B}(H),\ \|f\|_\infty\leq 1} \left|\int f(x)\lambda(dx)\right|. \tag{4.44}$$

Since $\|L_t - L_t\circ\theta_1\|_{var} \leq \frac{2}{t}$, we obtain for $t > 2/\delta$,

$$\begin{aligned}\mathbb{P}_\nu \left(L_t \in F\right) &\leq \mathbb{P}_\nu \left(L_t\circ\theta_1 \in F_\delta\right)\\ &= \mathbb{P}_{\tilde{\nu}} \left(L_t \in F_\delta\right)\\ &= \mathbb{P}_{\tilde{\nu}} \left(L_t \in F_\delta \cap M_1(D(A^\alpha))\right)\end{aligned}$$

by the regularity properties of the solution under $\tilde{\nu}$, defined by Lemma 4.6.1.

Let us fix $0 < b < a$. Since $[J \leq b]$ is contained in the open F^c (the complement of F), for each $\nu_i \in [J\leq b]$, we can take a neighborhood $N(\nu_i,\delta_i)$ of ν_i included in F^c. Moreover $N(\nu_i,\delta_i)$ can be chosen of form

$$N(\nu_i,\delta_i) := \left\{\beta \in M_1(H),\ |\nu_i(f_{i,j}) - \beta(f_{i,j})| < \delta_i,\ \forall j = 1\ldots d_i\right\}$$

for a finite number d_i of bounded and measurable $f_{i,j}$ with $\|f_{i,j}\|_\infty \leq 1$ for $1\leq j\leq d_i$. In particular $F \subset N(\nu_i,\delta_i)^c$.

Now, by Lemma 4.5.1, for any $b < a$, $[J\leq b]$ is compact in $(M_1(D(A^\alpha)),\tau)$ and so in $(M_1(H),\tau)$ since $M_1(D(A^\alpha))$ is just a borelian subset of $M_1(H)$. So, we can extract a finite number p of $\nu_i \in [J\leq b]$ such that

$$[J\leq b] \ \subset\ \cup_{i=1}^{i=p} N(\nu_i,\delta_i/2) \ \subset\ \cup_{i=1}^{i=p} N(\nu_i,\delta_i) \ \subset\ F^c.$$

We now prove that if $\delta \leq \min_{i=1\ldots p}\delta_i/2$, then

$$\cup_{i=1}^{i=p} N(\nu_i,\delta_i/2) \subset {F_\delta}^c. \tag{4.45}$$

Indeed, if $\nu \in F_\delta$ we can find $\beta \in F$ such that $\|\nu - \beta\|_{var} \leq \delta$. For any $i = 1 \dots p$, since $F \subset N(\nu_i, \delta_i)^c$, there is some j such that

$$|\beta(f_{i,j}) - \nu_i(f_{i,j})| \geq \delta_i.$$

With (4.44) and the fact that $\|f_{i,j}\|_\infty \leq 1$, we obtain

$$\begin{array}{rcl} |\nu(f_{i,j}) - \nu_i(f_{i,j})| & \geq & |\beta(f_{i,j}) - \nu_i(f_{i,j})| - |\beta(f_{i,j}) - \nu(f_{i,j})| \\ & \geq & \delta_i - \delta \\ & \geq & \delta_i/2 \end{array}$$

for $\delta \leq \min_{i=1\dots p} \delta_i/2$. So if $\nu \in F_\delta$, then $\nu \in N(\nu_i, \delta_i/2)^c$, for any $i = 1 \dots p$ and (4.45) is satisfied.

We obtain for $C > 0$ as in Lemma 4.6.1 and by the upper bound in Lemma 4.5.1

$$\begin{array}{rl} & \limsup_{T\to\infty} \dfrac{1}{T} \log \sup_{\nu \in \mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu (L_t \in F) \\ \leq & \limsup_{T\to\infty} \dfrac{1}{T} \log \sup_{\tilde{\nu} \in \mathcal{M}^*_{\lambda_0,\exp(\lambda_0 C)L}} \mathbb{P}_{\tilde{\nu}} (L_t \in F_\delta) \\ \leq & \limsup_{T\to\infty} \dfrac{1}{T} \log \sup_{\tilde{\nu} \in \mathcal{M}^*_{\lambda_0,\exp(\lambda_0 C)L}} \mathbb{P}_{\tilde{\nu}} \left(L_t \in \cap_{i=1\dots p} N(\nu_i, \delta_i/2)^c\right) \\ \leq & - \inf_{\nu \,\in\, \cap_{i=1}^{i=p} N(\nu_i,\delta_i/2)^c} J(\nu) \\ \leq & -b \end{array}$$

since the closed subset $\cap_{i=1}^{i=p} N(\nu_i, \delta_i/2)^c$ is contained in $[J \leq b]^c$. Noting that $0 < b < a$ is arbitrary, we obtain the upper bound (iii) in Theorem 4.1.1, which finishes the proof of Theorem 4.1.1. □

4.7 Extension to unbounded functionals

Let us now precise how the strong τ topology in Theorem 4.1.1, and the exponential estimate (4.23) imply Proposition 4.1.3 and more generally Corollary 4.1.2. In the sequel we suppose that our assumption (4.6) is satisfied for some $\frac{1}{4} < \alpha < \frac{1}{2}$, and that

$$0 < \lambda_0 < \frac{\lambda_1}{2\|Q\|}$$

is a fixed real number.

Proof of Corollary 4.1.2. For a measurable function $f : D\left(A^{\frac{1}{2}}\right) \to \mathbb{B}$, bounded on balls, let us consider $f_n : H \to \mathbb{B}$,

$$f_n(x) := \begin{cases} f(x) & \text{if } x \in D\left(A^{\frac{1}{2}}\right),\ |A^{\frac{1}{2}} x| \leq n \\ 0 & \text{otherwise} \end{cases} \tag{4.46}$$

which is far from being continuous, but is measurable and bounded on H. Since $\nu \to \nu(f_n)$ is continuous from $M_1(H)$ to $\mathbb{B}$ by [9, Lemma 3.3.8], $L_T(f_n)$ satisfies the LDP on $\mathbb{B}$ by Theorem 4.1.1 and a standard contraction principle.

Now, by the approximation lemma in large deviations (see [9, Lemma 2.1.4]), it remains to prove for any $L > 0$,

$$\limsup_{n\to\infty} \sup_{\beta\,:\,J(\beta)\le L} \|\beta(f_n) - \beta(f)\|_{\mathbb{B}} = 0 \tag{4.47}$$

and for any $\delta > 0$

$$\lim_{n\to\infty} \limsup_{T\to\infty} \frac{1}{T} \log \sup_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu\left(\|L_T(f - f_n)\|_{\mathbb{B}} > \delta\right) = -\infty. \tag{4.48}$$

Thanks to our condition (4.9) on f, we can construct a sequence $(\varepsilon(n))_n$ decreasing to 0 such that, once $|A^{\frac{1}{2}}x| \ge n$, we have

$$\|f(x)\|_{\mathbb{B}} \le \varepsilon(n)|A^{\frac{1}{2}}x|^2.$$

Denoting by 1_Γ the characteristic function of the set Γ, we have for any β satisfying $J(\beta) < L$,

$$\begin{aligned}
\|\beta(f_n) - \beta(f)\|_{\mathbb{B}} &= \left\|\beta\left(f 1_{\{|A^{\frac{1}{2}}x|\ge n\}}\right)\right\|_{\mathbb{B}} \\
&\le \beta\left(\varepsilon(n)|A^{\frac{1}{2}}x|^2 1_{\{|A^{\frac{1}{2}}x|\ge n\}}\right) \\
&\le \frac{\varepsilon(n)}{\lambda_0}\beta\left(\lambda_0|A^{\frac{1}{2}}x|^2\right) \\
&\le \frac{\varepsilon(n)}{\lambda_0}\left(L + \lambda_0\left(tr(Q) + \frac{\lambda_1|f|^2}{\lambda_1 - 2\|Q\|\lambda_0}\right)\right)
\end{aligned}$$

by using (4.25). Thus (4.47) follows. Let us also evaluate

$$\begin{aligned}
\mathbb{P}_\nu\left(\|L_T(f - f_n)\|_{\mathbb{B}} > \delta\right) &= \mathbb{P}_\nu\left(\left\|\frac{1}{T}\int_0^T f(X_s) - f_n(X_s)ds\right\|_{\mathbb{B}} > \delta\right) \\
&\le \mathbb{P}_\nu\left(\frac{1}{T}\int_0^T \varepsilon(n)|A^{\frac{1}{2}}X(s)|^2 1_{\{|A^{\frac{1}{2}}X(s)|\ge n\}} ds > \delta\right) \\
&\le \mathbb{P}_\nu\left(\int_0^T \lambda_0|A^{\frac{1}{2}}X(s)|^2 1_{\{|A^{\frac{1}{2}}X(s)|\ge n\}} ds > \frac{\lambda_0 T\delta}{\varepsilon(n)}\right) \\
&\le \exp\left(-\frac{\lambda_0 T\delta}{\varepsilon(n)}\right)\mathbb{E}^\nu \exp\left(\lambda_0\int_0^T |A^{\frac{1}{2}}X(s)|^2 ds\right)
\end{aligned}$$

so that (4.48) is consequence of (4.23). $\square$

Proof of Proposition 4.1.3. It is a particular case of Corollary 4.1.2, since the choice $f(x) = x$ on $\mathbb{B} := D\left(A^{\frac{1}{2}}\right)$ is allowed. $\square$

Acknowledgements :

The author wishes to thank Professor Liming Wu for many helpful suggestions and discussions.

Bibliography

[1] A. Amirdjanova, J. Xiong, Large deviation principle for a stochastic Navier-Stokes equation in its vorticity form for a two-dimensional incompressible flow, *Discrete. Contin. Dyn. Syst. Ser. B,* **6***, 651-666 (2006).*

[2] A. Bensoussan, R. Temam, Equations Stochastiques du Type Navier-Stokes , *J. Funct. Anal.,* **13***, 195-222 (1973).*

[3] M. Capinski, D. Gatarek, Stochastic Equations in Hilbert Space with application to Navier-Stokes Equations in Any Dimension, *J. Funct. Anal.,* **126***, 26-35 (1994).*

[4] M.H. Chang, Large deviation for Navier-Stokes equations with small stochastic perturbation, *Appl. Math. and Comp.,* **76***, 65-93 (1996).*

[5] R. Collina, R. Livi, A. Mazzino, Large deviation approach to the randomly forced Navier-Stokes equation, *J. Stat. Phys.,* **118***, 451-479 (2005).*

[6] G. Da Prato and A. Debussche, Ergodicity for the 3D stochastic Navier-Stokes equations, *J. Math. Pures. Appl,* **82***, 877-947 (2003).*

[7] G. Da Prato and J. Zabczyk, *Ergodicity for Infinite Dimensional Systems*, Cambridge University Press (1996).

[8] A. Dembo and O. Zeitouni, *Large Deviations Techniques and applications, 2nd edition*, Springer-Verlag (1998).

[9] J.D. Deuschel and D.W. Stroock, *Large Deviations*, Pure Appl. Math., Vol. 137, Academic Press, San Diego (1989).

[10] M.D Donsker and S.R.S Varadhan, Asymptotic evalutation of certain Markov process expectations for large time, I-IV, *Comm. Pur. Appli. Math,* **28**, *1-47 and 279-301 (1975);* **29**, *389-461 (1976);* **36**, *183-212 (1983).*

[11] Weinan E, J.C. Mattingly, Y.G. Sinai, Gibbsian dynamics and ergodicity for the stochastic forced Navier-Stokes equation, *Comm. Math. Phys.,* **224**, *83-106 (2002).*

[12] B. Ferrario, Ergodic results for stochastic Navier-Stokes Equations, *Stochastics. Stoch. Rep.,* **60**, *271-288 (1997).*

[13] F. Flandoli, Dissipativity and invariant measures for stochastic Navier-Stokes Equations, *NoDEA,* **1**, *403-428 (1994).*

[14] F. Flandoli, D. Gatarek, Martingale and stationary solutions for stochastic Navier-Stokes Equations, *Proba. Theo. Rel. Fields,* **102**, *367-391 (1995).*

[15] F. Flandoli, B. Maslowski, Ergodicity of the 2D Navier-Stokes Equation Under Random Perturbation, *Communications in Mathematical Physics,* **171**, *119-141 (1995).*

[16] C. Foias, O. Manley, R. Rosa and R. Temam *Navier-Stokes Equations and Turbulence*, Cambridge University Press (2001).

[17] B. Goldys, B. Maslowski, Exponential ergodicity for stochastic Burgers and 2D Navier-Stokes equation, *J. Funct. Anal.,* **226**, *Nr 1, 230-255 (2005).*

[18] M. Gourcy, Large Deviation Principle of Occupation Measure for Stochastic Burgers Equation, *preprint, submitted to annals of I.H.P (2006).*

[19] M. Gourcy, *Inégalités Log-Sobolev pour la loi d'une diffusion et Grandes déviations pour des EDP stochastiques*, PHD Thesis (2006), Université Blaise Pascal.

[20] M. Hairer, J.C. Mattingly, E. Pardoux, Malliavin calculus for highly degenerate 2D stochastic Navier-Stokes equations, *C. R. Math. Acad. Sci. Paris* **339**, *(11), 793–796 (2004).*

[21] M. Hairer, J.C. Mattingly, Spectral Gaps in Wasserstein distances and the 2D Navier-Stokes Equations, *preprint, (2006).*

[22] S. Kuksin, A. Shirikyan, Stochastic Dissipative PDE's and Gibbs Measures, *Comm. Math. Phys.,* **213**, *291-330 (2000).*

[23] J.C. Mattingly, Exponential convergence for the Stochastically forced Navier-Stokes equations and other partially dissipative dynamics, *Comm. Math. Phys.,* **230***, (3), 421-462 (2002).*

[24] C. Odasso, *Méthodes de couplage pour des équations stochastiques du type Navier-Stokes et Schrödinger*, PHD thesis (2005), Université de Rennes 1.

[25] S.S. Sritharan, P. Sundar, Large Deviations for the Two-dimensional Navier-Stokes Equation with Multiplicative Noise, *to appear in S.P.A (2006).*

[26] R. Temam, *Navier-Stokes Equations*, North-Holland (1977).

[27] L. Wu, Uniformly integrable operators and large deviations for Markov processes, *J.Funct.Anal.* **172***, 301-376 (2000).*

[28] L. Wu, Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems, *S.P.A.,* **91***, 205-238 (2001).*

Chapitre 5

Existence of an invariant measure for a stochastically kicked Navier-Stokes equation

Chapitre rédigé en anglais.

Abstract : We establish a new version of the drift criterion for the existence of an invariant measure well adapted to the case of an infinite dimensional state space valued Markov chain. Applications to a Stochastic Navier-Stokes equation are given. we only require a noise with logarithmic integrable norm.

Keywords : Stochastic Navier-Stokes Equation, Markov Chain, Invariant Measure.

AMS Subject Classification 2000: 60J27, 76D05.

5.1 Invariant measure for the stochastically kicked 2D Navier-Stokes equations

Let us introduce the two-dimensional Navier-Stokes equations. The incompressible Navier-Stokes equations describe the evolution of an incompressible fluid. They are most frequently written in terms of the velocity field at each point in the domain U. We restrict our discussion to the case of two spatial dimensions.

$$\begin{cases} \frac{du}{dt} + (u.\nabla u) + \nabla p = \nu \Delta u + g \\ \text{div } u = 0 \\ u(0,\xi) = u_0(\xi) \text{ and } \int_U u(t,\xi) d\xi = 0 \text{ for } t \geq 0 \end{cases} \tag{5.1}$$

Here, we restrict also to the 2Π-periodic case (with mean zero flow). On a periodic domain of $\mathbb{T}^2 = [0, 2\Pi] \times [0, 2\Pi]$, the standard Fourier basis form a useful eigenbasis for $(-\Delta)$. In the usual way, we exclude the pressure p from the equations, applying to (5.1) the projection to the linear space of divergence free vector fields. To this end, $\mathbb{L}^2$ will denote the closure in the L^2 topology, of divergence free, mean zero, $\mathcal{C}^\infty$ vector fields on the two dimensional torus $\mathbb{T}^2$.

Define $B(u, v) = P_{div}(u.\nabla)v$ and $Au = -P_{div}\Delta u$ where P_{div} is the L^2 projection operator onto the space of divergence-free vector field. We obtain the equation

$$\begin{cases} du(t,\xi) + \nu Au(t,\xi)dt + B(u,u)dt + = f dt \\ u(0,\xi) = u_0(\xi) \in \mathbb{L}^2 \text{ and } \int_{\mathbb{T}^2} u(t,\xi) d\xi = 0 \text{ for } t \geq 0 \end{cases} \tag{5.2}$$

where $f = P_{div} g$.

Remark that, with this setting, an eigenbasis for A in $\mathbb{L}^2$ is given by

$$\left\{ e_k(\xi) = (-k_2, k_1) \frac{e^{ik.\xi}}{|k|}, \ k = (k_1, k_2) \in \mathbb{Z}^2 \right\}$$

where $\xi = (\xi_1, \xi_2) \in \mathbb{T}^2$, $|k|^2 = k_1^2 + k_2^2$, and the corresponding eigenvalues are $\lambda_k = |k|^2$.

In mathematical literature, it is common to assume the random force $P_{div} f(t, x)$ to be random fields that are smooth in x, while as a function of time t they are white noises, see [1] and [2]. In this note, we choose another popular mathematical model, namely a *kick force model*, already studied by Kuksin and Shirikyan [4, 5, 6] or Masmoudy and Young [8].

A kick force correspond to the situation when the system gets smooth random kicks with some time period T and evolves freely between the kicks. It means that the k^{th} kicks change a solution $u(kT, \xi)$ to $u(kT + 0, \xi) = u(kT, \xi) + \eta_k(\xi)$, while between the kicks, $u(t, x)$ satisfies the equation (5.2) with $g = 0$.

This model is described by equation (5.2) in which the force is a δ-function of time t:

$$f = P_{div} g = \sum_{k \geq 0} \delta(t - kT) \eta_k(x).$$

Here the η_k are independent identically distributed smooth random fields with range in $\mathbb{L}^2$.

Due to the kick nature of the force, a solution $u(t,\xi)$ is completely described by its values $u_k(\xi)$ at the points $t = kT + 0,\ k \geq 0$:

$$u_k(\xi) := u(kT + 0, \xi),\ \ k \geq 0.$$

For simplicity we fix $T = 1$, but the result holds for any $T > 0$. From now on we treat (5.2) as a discrete time random dynamical system in $\mathbb{L}^2$:

$$u_k = S(u_{k-1}) + \eta_k, \tag{5.3}$$

where $S : \mathbb{L}^2 \to \mathbb{L}^2$ is a time-one shift along trajectories of solutions to equation (5.2) without force. Note that the map S is well defined acording to classical theory [10].

Our main result is

Theorem 5.1.1. *If for some $c_0 > 0$*

$$\mathbb{E} \log(1 + c_0 \|\eta_k\|_{\mathbb{L}^2}) < \infty \tag{5.4}$$

then the $\mathbb{L}^2$-valued Markov chain defined by (5.3) *admits an invariant measure.*

Fortunately, it includes the non forced case, where the dirac in 0 is the unique invariant measure (it is well known that without external forces the system relaxes to its equilibrium).

The existence of an invariant measure for this problem was first established by Kuksin and Shirikyan in [4] under more stringent assumptions on the noise than (5.4). In fact their aim was the uniqueness under some additional hypothesis. The exponential mixing property was later proved by Masmoudi and Young [8] and also by Kuksin and Shirikyan [6, 9].

In reality, we shall establish a very general criterion for the existence of invariant measure, which is not only well adapted to Navier-Stokes equation, but also to many other models.

5.2 General result in an abstract setting

In this section, we leave the Navier-Stokes equations for one moment, and we work in an abstract setting to illustrate the generality of our result.

5.2.1 Notations and assumptions

Let (E, d) be a Polish space. In the sequel, $\mathbb{1}_A$ will be the characteristic function of a subset $A \subset E$ and

$$B(R) = \{x \in E \text{ such that } d(0, x) \leq R\}.$$

Let $P(x, dy)$ be a Markov kernel on E. It can be regarded as an operator acting on some Banach lattice of measurable functions on E via $Pf(x) := \int_E f(y)P(x, dy)$, but also as an operator on $M_1(E)$, the set of probability measures on E, via $mP(A) = \int_E P(x, A)m(dx)$. We recall that P is called (weak) Feller if the set of bounded continuous functions is stable under the preceding operator. A measure μ on E is invariant for P if $\mu(A) = \int_E P(x, A)\mu(dx)$, that is if $\mu P = \mu$. The invariant measures contain all information concerning the stability of the system. In practice, we restrict ourself to the ergodic measures which are the extremal points of the convex set of invariant measures. Thus, existence of an invariant measure implies existence of an ergodic invariant measure (see [1]).

When we try to apply traditional criteria for the existence as those exposed in the book of Meyn and Tweedie [7], we need conditions of return to compact subsets. In the infinite dimensional setting, we do not know an efficient characterization of such compact sets (excepted Rellich theorem for Sobolev spaces). Our goal is to find simple conditions on P under which existence of an invariant measure is guaranteed.

We will say that the Markov kernel P satisfies the mean drift condition if the following statement is fulfilled.

Assumption 5.2.1. *There exists a continuous and locally bounded (i.e. bounded over bounded subsets) function $u : E \to [0, \infty[$ and two real numbers $R > 0$ and $b > 0$ such that*

$$\forall x \in E, \qquad Pu(x) \leq u(x) - 1 + b\mathbb{1}_{B(R)}(x). \tag{5.5}$$

In fact we can replace (5.5) by $Pu(x) \leq u(x) - c + b\mathbb{1}_{B(R)}(x)$ where $c > 0$ is any fixed real number. Recently, Wu [11] gives estimation of the essential radius spectral of P by means of the data given by Assumption 5.2.1.

In a classical way, $B(R)$ will be a kind of *recurrent* set for our Markov chain. One seeks now an invariant measure for the sampled chain based on the return time in the set $B(R)$.

Assumption 5.2.2. *The kernel P is Feller and for any $R > 0$, the family of probability measures $\{P(x, .), x \in B(R)\}$ is tight.*

We recall that the above assumption implies the existence for any $\varepsilon > 0$ of a compact subset K_ε of E satisfying $P(x, K_\varepsilon) > 1 - \varepsilon$ for all x in $B(R)$.

5.2.2 Results

In this section we establish the

Theorem 5.2.3. *Any Markov kernel $P(x, dy)$ satisfying Assumption 5.2.1 and Assumption 5.2.2 admits an invariant measure.*

In the sequel $\Phi := \{X_k\}_{k\in\mathbb{N}}$ will be the homogenous Markov chain with $P(x,dy)$ as one step transition function, $\mathbb{E}^x$ is the expectation with initial value $X_0 = x$ for the process, and $\mathbb{E}^m$ the expectation with random initial value X_0 of law m.

Let us sketch the proof of this theorem. First, we check that the return time $\tau_{B(R)}$ in $B(R)$ is integrable.

Lemma 5.2.4. *Let $\tau_{B(R)} := \inf\{n \geq 1 \ s.t\ X_n \in B(R)\}$. For $R, b > 0$ given by Assumption 5.2.1, we have*

$$\mathbb{E}^x \tau_{B(R)} \leq u(x) + b\mathbb{1}_{B(R)}(x), \qquad \forall x \in E$$

Then, we consider the Markov chain composed by the successive returns in $B(R)$, and we establish the

Lemma 5.2.5. *Under Assumption 5.2.1 and Assumption 5.2.2, the Markov kernel defined by $Q(x,dy) := \mathbb{P}_x\left(X_{\tau_{B(R)}} \in dy\right)$ admits an invariant measure m.*

In that case, it is well known and suggested by the ergodic theorem that an invariant probability measure μ for $P(x,dy)$ is given by the mean time spent in set between two visits in the ball $B(R)$. That is

$$\mu(A) := \frac{\mathbb{E}^m\left(\sum_{k=1}^{\tau_{B(R)}} \mathbb{1}_A(X_k)\right)}{\mathbb{E}^m \tau_{B(R)}}. \tag{5.6}$$

5.2.3 Proofs

Let us prove the two preceding lemmas and our new criterion.

Proof of Lemma 5.2.4

Let us evaluate

$$\begin{aligned} u\left(X_{n\wedge\tau_{B(R)}}\right) - u(x) &= \sum_{k=1}^{n\wedge\tau_{B(R)}} u(X_k) - u(X_{k-1}) \\ &= \sum_{k=1}^{n\wedge\tau_{B(R)}} u(X_k) - Pu(X_{k-1}) + \sum_{k=1}^{n\wedge\tau_{B(R)}} (Pu-u)(X_{k-1}) \\ &= M_{n\wedge\tau_{B(R)}} + \sum_{k=0}^{\left(n\wedge\tau_{B(R)}\right)-1} (Pu-u)(X_k) \end{aligned}$$

where $M_n := \sum_{k=1}^n u(X_k) - Pu(X_{k-1})$ is a martingale. Then with Assumption 5.2.1, we obtain

$$u\left(X_{n\wedge\tau_{B(R)}}\right) - u(x) \leq M_{n\wedge\tau_{B(R)}} + \sum_{k=0}^{\left(n\wedge\tau_{B(R)}\right)-1} \left(-1 + b\mathbb{1}_{B(R)}(X_k)\right).$$

But if $n \wedge \tau_{B(R)} \geq 2$, it is clear that $X_k \notin B(R)$ for $1 \leq k \leq \left(n \wedge \tau_{B(R)}\right) - 1$. Hence, we have

$$\begin{aligned} n \wedge \tau_{B(R)} &\leq u(x) - u\left(X_{n\wedge\tau_{B(R)}}\right) + M_{n\wedge\tau_{B(R)}} + b\mathbb{1}_{B(R)}(x) \\ &\leq u(x) + M_{n\wedge\tau_{B(R)}} + b\mathbb{1}_{B(R)}(x). \end{aligned}$$

Taking expectation, the monotone convergence allows us to conclude. □

Proof of Lemma 5.2.5

We proceed in three steps.

- step 1: For all n in $\mathbb{N}^*$ and for all $\varepsilon > 0$, there exists a compact set C in E such that

$$\inf_{x\in B(R)} \mathbb{P}_x \left(X_k \in C\ , \forall\ k,\ 1 \leq k \leq n\right) \geq 1 - \varepsilon$$

To prove that, we use an argument based on $\beta_w(P)$ the measure of non-compactness (for the weak topology) of a transition semigroup P, introduced by Wu [11]. Let us recall, the definition

$$\beta_w(P) := \inf_{K\subset\subset E} \sup_{x\in E} P(x, K^c)$$

where P is a bounded nonnegative kernel on E, and the notation $K \subset\subset E$ means that K is a compact subset of E. It is clear that $\beta_w(P) = 0$, if and only if the set of measure $\{P(x, dy), x \in E\}$ is tight. We recall also a useful result proved by Wu ([11] p 263).

Lemma 5.2.6. *[11] Assume that*

$$\textit{for all compact subset } K \textit{ of } E, \qquad \beta_w(\mathbb{1}_K P) = 0. \tag{5.7}$$

Then for any nonnegative bounded kernel Q

$$\beta_w(QP) \leq \beta_w(Q)\beta_w(P).$$

Within our framework (5.7) is clearly implied by the Feller property of P which gives exactly the continuity of the application $x \in E \mapsto P(x, dy) \in M_1(E)$ equipped with the weak topology.

Fix $n \in \mathbb{N}^*$ and $\varepsilon > 0$. For $1 \leq k \leq n$, Lemma 5.2.6 implies

$$\beta_w(\mathbb{1}_{B(R)}P^k) \leq \beta_w(\mathbb{1}_{B(R)}P)\beta_w(P^{k-1}) = 0$$

by Assumption 5.2.2. So we can find a compact subset C_k of E such that

$$\forall x \in B(R),\ \ P^k(x, C_k^c) \leq \frac{\varepsilon}{n}.$$

Now, let us consider the compact $C = \cup_{k=1}^{n} C_k$. It satisfies for all x in $B(R)$,

$$\begin{array}{rcl} \mathbb{P}_x \left(X_k \in C\ , \forall\ k,\ 1 \leq k \leq n \right) & = & 1 - \mathbb{P}_x \left(\exists\ k,\ 1 \leq k \leq n,\ X_k \in C^c \right) \\ & \geq & 1 - \displaystyle\sum_{k=1}^{n} \mathbb{P}_x \left(X_k \in C_k^c \right) \\ & \geq & 1 - \varepsilon. \end{array}$$

- step 2: $\{Q(x,dy), x \in B(R)\}$ is tight.
  By Lemma 5.2.4,
  $$\mathbb{E}^x(\tau_{B(R)}) \leq u(x) + b\mathbb{1}_{B(R)}(x)$$
  where the right hand side is bounded in $B(R)$ according to Assumption 5.2.1. So, for any $\varepsilon > 0$, we can find an integer n, independent of x in $B(R)$, such that
  $$\forall x \in B(R), \quad \mathbb{P}_x \left(\tau_{B(R)} > n \right) < \varepsilon.$$
  With this choice for n in the first step, and the corresponding compact C, which is also compact for the inherited topology on $B(R)$, we obtain for all $x \in B(R)$,
  $$\begin{array}{rcl} Q(x,C) & = & \mathbb{P}_x \left(X_{\tau_{B(R)}} \in C \right) \\ & \geq & \mathbb{P}_x \left(X_{\tau_{B(R)}} \in C\ ,\ \tau_{B(R)} \leq n \right) \\ & \geq & \mathbb{P}_x \left(X_k \in C,\ \forall k,\ 1 \leq k \leq n\ ,\ \tau_{B(R)} \leq n \right) \\ & \geq & \mathbb{P}_x \left(X_k \in C,\ \forall k,\ 1 \leq k \leq n \right) - \mathbb{P}_x \left(\tau_{B(R)} > n \right) \\ & \geq & 1 - 2\varepsilon \end{array}$$
  using the elementary fact that $\mathbb{P}(A \cap B) \geq \mathbb{P}(A) - \mathbb{P}(B^c)$.

- step 3: According to Theorem 12.1.2 in [7] for proving that $Q(x,dy)$ admits at least one invariant probability m, it is sufficient to show that $Q(x,dy)$ is bounded in probability on average. It means that, for any x in $B(R)$, the set of probability measures $\{\bar{Q}_n(x,dy),\ n \in \mathbb{N}^*\}$ is tight, where
  $$\bar{Q}_n(x,dy) := \frac{1}{n} \sum_{k=1}^{n} Q^k(x,dy).$$
  Let $\varepsilon > 0$ and the compact C as given by the preceding step 1, we obtain for any x in $B(R)$,
  $$Q^k(x, C^c) \ \ \leq \ \ \sup_{y \in B(R)} Q(y, C^c)$$

$$\begin{aligned} &\leq \sup_{y\in B(R)} \mathbb{P}_y\left(X_{\tau_{B(R)}} \in C^c\right) \\ &\leq 2\varepsilon \end{aligned}$$

So, $\bar{Q}_n$ is tight and $Q(x,dy)$ generate a Feller Markov chain bounded in probability. □

Proof of Theorem 5.2.3

Let us check that the measure μ defined by (5.6) satisfies $\mu P{=}\mu$. For any function $f \in C_b(E)$, we know by the Feller property of P that Pf is also a bounded and continuous function on E. So, Lemma 5.2.5 implies that

$$\int_{B(R)} Pf(x)m(dx) = \int_{B(R)}\int_{B(R)} Pf(y)Q(x,dy)m(dx)$$

i.e $\mathbb{E}^m Pf = \mathbb{E}^m Pf(X_{\tau_{B(R)}})$ because m is an invariant measure for the Markov kernel on $B(R)$ defined by $Q(x,dy) := \mathbb{P}_x\left(X_{\tau_{B(R)}} \in dy\right)$. Then,

$$\begin{aligned} \mathbb{E}^m\left(\sum_{k=1}^{\tau_{B(R)}} f(X_k)\right) &= \mathbb{E}^m\left(\sum_{k=1}^{\infty} f(X_k)\mathbb{1}_{k\leq\tau_{B(R)}}\right) \\ &= \mathbb{E}^m\left(\sum_{k=1}^{\infty} Pf(X_{k-1})\mathbb{1}_{k\leq\tau_{B(R)}}\right) \\ &= \mathbb{E}^m\left(\sum_{k=1}^{\tau_{B(R)}-1} Pf(X_k)\right) + \mathbb{E}^m\left(Pf\right) \\ &= \mathbb{E}^m\left(\sum_{k=1}^{\tau_{B(R)}} Pf(X_k)\right). \end{aligned}$$

So, for any bounded and continuous function f on E, we have

$$\begin{aligned} \mu(f) &= \frac{\mathbb{E}^m\left(\sum_{k=1}^{\tau_{B(R)}} Pf(X_k)\right)}{\mathbb{E}^m\tau_{B(R)}} \\ &= \mu(Pf). \end{aligned}$$

where μ is given by (5.6). □

5.2.4 Application to a discrete dynamical system perturbed simple kicks

Let us consider the discrete dynamical system on a Banach space E defined by

$$X_{n+1} = S(X_n) + W_{n+1}\ ,\ X_0 \in E,\ n \geq 0 \tag{5.8}$$

where $(W_n)_{n\geq 1}$ is a sequence of E-valued independent identically distributed random variables, with the same law as, say W. We denote by $|x|$ the norm of an element x in E.

We assume the following

Assumption 5.2.7. *The data defining (5.8) are such that:*

- *(a) the operator $S : E \to E$ is continuous and for any $R > 0$, $S(B(R))$ is relatively compact,*
- *(b) there exists $0 < r < 1$ and $R_0 > 0$ such that $|S(x)| \leq r|x|, \ \forall |x| \geq R_0$,*
- *(c) we can find $c_0 > 0$ such that $\mathbb{E}\log(1 + c_0|W|) < \infty$.*

Now, we establish as a corollary of Theorem 5.2.3 the

Proposition 5.2.8. *Under Assumption 5.2.7, the dynamical system defined by (5.8) admits an invariant measure.*

Proof. We check Assumptions 5.2.1 and 5.2.2 for the transition kernel of (5.8).

Let r, R_0 and c_0 be given by Assumption 5.2.7. Fix $\varepsilon > 0$ such that $r(1+\varepsilon) < 1$. Then, by dominated convergence theorem, we can find c such that $0 < c < c_0$ and

$$\mathbb{E}\log(1 + c|W|) + \log\left(r(1+\varepsilon)\right) := -\delta < 0. \tag{5.9}$$

Consider the locally bounded test function $v(x) = \log(1 + c|x|)$. Since, we have

$$Pv(x) = \mathbb{E}\log\left(1 + c|S(x) + W|\right)$$

Assumption 5.2.7 implies clearly that $Pv - v$ is bounded on $B(R)$ for any $R > 0$. Hence, for establishing Assumption 5.2.1, it is sufficient to find $R > 0$ such that,

$$\forall x \notin B(R), \quad Pv(x) \leq -1 + v(x). \tag{5.10}$$

Recall that for any $a, b \geq 0$,

$$\log(1 + a + b) \leq \log(1 + a) + \log(1 + b)$$

we remark that for $|x| \geq R_0$

$$\begin{array}{rcl} Pv(x) & = & \mathbb{E}\log\left(1 + c|S(x) + W|\right) \\ & \leq & \log\left(1 + c|S(x)|\right) + \mathbb{E}\log\left(1 + c|W|\right) \\ & \leq & \log\left(1 + cr|x|\right) + \mathbb{E}\log\left(1 + c|W|\right). \end{array}$$

Now for r, ε as above, we can choose $R_\varepsilon := \frac{1-r(1+\varepsilon)}{r\varepsilon} > 0$ such that

$$\forall y \geq R_\varepsilon, \qquad \log(1 + ry) \leq \log(1 + y) + \log\left(r(1+\varepsilon)\right).$$

Using (5.9) and the above inequality for $y = c|x|$, we obtain for all $|x| \geq R_0 \wedge R_\varepsilon/c$,

$$\begin{array}{rcl} Pv(x) & \leq & \log(1 + cr|x|) + \mathbb{E}\log(1 + c|W|) \\ & \leq & \log(1 + c|x|) + \log(r(1+\varepsilon)) + \mathbb{E}\log(1 + c|W|) \\ & \leq & v(x) - \delta \end{array}$$

Finally, for $R := R_0 \wedge R_\varepsilon/c$ in (5.10), the drift assumption 5.2.1 is satisfied with the test function $u(x) = v(x)/\delta$.

It is also clear that (a) implies Assumption 5.2.2, for

$$\{P(x, dy)\ ;\ x \in B(R)\} \subset \{\text{law of } y + W;\ y \in S\,(B(R))\}$$

which is tight. The property follows by Theorem 5.2.3. □

Remarks 5.2.9. *(i) For example, Assumption 5.2.7 (c) is clearly satisfied if the noise is such that*

$$\mathbb{E}|W| < \infty.$$

(ii) Sometimes a little dissipative force is sufficient: if the condition (c) is replaced by the more stringent $\mathbb{E}|W| \leq C_0$, then (b) can be replaced by the weaker (b') there exists $C > C_0$ and $R_0 > 0$ such that $|S(x)| \leq |x| - C$, $\forall |x| \geq R_0$. For checking the result, consider $v(x) = 1 + |x|$.

(iii) In the case when E is finite dimensional, we do not need the compactness of S, and the condition become very simple to handle. For example, let us fix $E = \mathbb{R}$ and $u(x) = x^2$. Then, for $S(x) = x - \log(x)$, the criterion is satisfied for any noise verifying $\mathbb{E}|W|^2 < \infty$, whereas for $S(x) = x - sgn(x)\alpha$, it is only satisfied by noise verifying $\mathbb{E}|W|^2 < \infty$ and $\mathbb{E}|W| < \alpha$.

5.3 Application to the NS equations

According to the preceding section, for proving Theorem 5.1.1, we must check parts (a), (b) and (c) of Assumption 5.2.7 for the Markov chain (5.3). Note that (c) is our basic condition on the noise.

For (b), let us recall classical properties of the Navier-Stokes equations on the 2-D torus (see for instance the famous book of Temam [10], p196 or the book of Galdy [3]). Indeed, we have the following energy estimate

$$\|u(t)\|_{L^2}^2 + 2\nu \int_0^t \|\nabla u(s)\|_{L^2}^2 ds = \|u_0\|_{L^2}^2, \quad \forall t \geq 0 \tag{5.11}$$

and the Poincaré inequality

$$\|u(t)\|_{L^2}^2 \leq \|\nabla u(t)\|_{L^2}^2, \quad \forall t \geq 0. \tag{5.12}$$

From (5.11) and (5.12), it follows that the function $t \to e^{2\nu t}\|u(t)\|_{L^2}^2$ is non increasing, and we obtain

$$\|S(u)\|_{L^2} \leq e^{-\nu}\|u\|_{L^2}.$$

That implies (b).

Finally (a) is a consequence of the following inequality established in the paper of Masmoudi and Young (estimate (15) in [8]):

$$\|S(u_0) - S(v_0)\|_{H^1} \leq C\|u_0 - v_0\|_{L^2},$$

recalling that the embedding of H^1 into L^2 is compact.

Bibliography

[1] G. Da Prato and J. Zabczyk, *Ergodicity for Infinite Dimensional Systems*, Cambridge University Press (1996).

[2] E. Weinan, J.C. Mattingly, Y.G. Sinai, Gibbsian dynamics and ergodicity for the stochastic forced Navier-Stokes equation, *Comm. Math. Phys.,* **224**, *(1) (2002).*

[3] G.P. Galdy. *An introduction to the mathematical theory of the Navier-Stokes equations, I et II*, Springer (1994).

[4] S. Kuksin, A. Shirikyan, Stochastic Dissipative PDE's and Gibbs Measures, *Comm. Math. Phys.,* **213**, *291-330 (2000).*

[5] S. Kuksin, A. Shirikyan, A Coupling approach to randomly forced non linear PDE's. I, *Comm. Math. Phys.,* **221**, *351-366 (2001).*

[6] S. Kuksin, A. Shirikyan, Ergodicity for the randomly forced 2D Navier-Stokes equations, *Math. Phys. Anal. Geom,* **4**, *147-195 (2001).*

[7] S.P Meyn and R.L. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag (1993).

[8] N. Masmoudi, L.S. Young, Ergodicity of the 2D-Navier-Stokes equations with Degenerate Stochastic Forcing, *Comm. Math. Phys.,* **227**, *461-481, (2002).*

[9] A. Shirikyan, Exponential mixing for 2D Navier-Stokes Equations Perturbed by an Unbounded Noise, *J. math. fluid mech.,* **6**, *169-193 (2004).*

[10] R. Temam, *Navier-Stokes equations*, North-Holland, (1984).

[11] L. Wu, Essential spectral radius for Markov semigroups (I): discrete time case *P.T.R.F.* **128**, *255-321 (2004).*

Chapitre 6

Large deviation principle of occupation measure for stochastic Burgers equation

Chapitre rédigé en anglais.

Article accepté aux Annales de l'Institut Henri Poincaré.

Abstract :

In this paper we obtain a Large Deviation Principle for the occupation measure of the solution to a stochastic Burgers equation which describes the exact rate of exponential convergence. This Markov process is strongly Feller and has a unique invariant measure. Moreover, the rate function is explicit: it is the Level-2 entropy of Donsker-Varadhan.

Résumé :

On obtient un Principe des Grandes Déviations pour la mesure d'occupation associée à la solution d'une équation de Burgers stochastique, décrivant la convergence exponentielle. La fonction de taux associée est l'entropie de niveau 2 de Donsker-Varadhan.

Keywords : Stochastic Burgers Equation, Large Deviations, Occupation Measure.

AMS Subject Classification 2000: 60F10, 60J35, 35Q53.

6.1 Introduction and Main results

Let $H = L^2(0,1)$ equipped with its norm $\|\cdot\|_2$. In this paper we are interested in the large time behavior of the solution to the following stochastic Burgers equation:

$$dX(t) = \left(\Delta X(t) + \frac{1}{2}D_\xi X^2(t)\right)dt + GdW(t)\ ;\ X(0,\xi) = x_0(\xi) \in H \tag{6.1}$$

where $G : H \to H$ is a bounded linear operator, $W(t)$ is a standard cylindrical Wiener process on H, and Δ is the Laplacian on $(0,1)$ with the Dirichlet boundary conditions. Indeed, the problem (6.1) is supplemented by:

$$X(t,0) = X(t,1) = 0, \quad t > 0.$$

It is well known that Δ is a negative, self-adjoint, non bounded operator on H with the domain of definition given by

$$D(\Delta) = \{u \in H^2(0,1) : u(0) = u(1) = 0\} = H^2(0,1) \cap H_0^1(0,1).$$

where

$$H_0^1 = \{x : [0,1] \to \mathbb{R};\ x \text{ is abs. continuous, } x(0) = x(1) = 0, \text{ and } \nabla x := D_\xi x \in H\}.$$

We assume that $tr(GG^*) < \infty$, i.e, the energy injected by the random force is finite, and that, for $Q = GG^*$,

$$Im\left((-\Delta)^{-\frac{\delta}{2}}\right) \subset Im\left(Q^{\frac{1}{2}}\right) \text{ for some } \frac{1}{2} < \delta < 1 \tag{6.2}$$

where $Im\left(Q^{\frac{1}{2}}\right)$ is the range of the operator $Q^{\frac{1}{2}}$. The last condition (6.2) means that the noise is not too degenerate. It is equivalent to say that the domain of definition of $(-\Delta)^{\frac{\delta}{2}}$ in H is contained in $Im\left(Q^{\frac{1}{2}}\right)$.

The above equation plays an important role in fluid dynamic for understanding of chaotic behavior. This stochastic model has been intensively studied for 10 years, in particular by Da Prato, Debussche, Dermoune, Weinan E, Gatarek, Khanin, Mazel, Sinai and Temam among many others (from a chronological point of view, see [7], [6], [11], [4], [5], [14]). About large deviations, small noise asymptotic was investigated by Cardon-Weber [3]. More recently, Goldys and Maslowsky proved the exponential ergodicity [16].

Let $M_1(H)$ (resp. $M_b(H)$) be the space of probability measures (resp. signed σ-additive measures of bounded variation) on H equipped with the Borel σ-field

$\mathcal{B}(H)$. The usual duality relation between $\nu \in M_b(H)$ and $f \in b\mathcal{B}(H)$, the set of bounded and measurable functions on H, will be denoted by

$$\nu(f) := \int_H f d\nu.$$

On $M_b(H)$ (or its subspace $M_1(H)$), we will consider the usual weak convergence topology $\sigma(M_b(H), C_b(H))$ and the so called τ-topology $\sigma(M_b(H), b\mathcal{B}(H))$, which is much stronger.

Our aim is to establish the large deviation principle (LDP in short) for the occupation measure L_t of the solution X (or empirical measure of level-2) given by

$$L_t(A) := \frac{1}{t}\int_0^t \delta_{X_s}(A)ds, \quad \forall A \in \mathcal{B}(H)$$

δ_a being the Dirac measure at a. Notice that L_t is an in $M_1(H)$-valued random variable. This is a traditional subject in probability since the pioneering work of Donsker and Varadhan [13]. The main innovation is that we deal about infinite dimensional diffusions for which their assumptions are not satisfied. For an introduction to large deviations we refer to the books of Deuschel-Stroock [12], and Dembo-Zeitouni [10].

Under (6.2), it is known that X_t is a Markov process with a unique invariant measure μ (cf [9]). So the ergodic theorem says that, almost surely under $\mathbb{P}_\mu$, L_t converges weakly to μ. We establish in this note a much more stronger result:

Theorem 6.1.1. *Assume that $tr(GG^*) < +\infty$ and (6.2) (throughout this paper). Let $0 < \lambda_0 < \frac{\pi^2}{2\|Q\|}$, where $\|Q\|$ is the norm of Q as an operator in H and*

$$\Phi(x) = e^{\lambda_0\|x\|_2^2}, \ \mathcal{M}_{\lambda_0,L} := \left\{\nu \in M_1(H) \ / \ \int_H \Phi(x)\nu(dx) \le L\right\} \tag{6.3}$$

The family $\mathbb{P}_\nu(L_T \in \cdot)$ as $T \to +\infty$ satisfies the large deviation principle (LDP) with respect to (w.r.t. in short) the topology τ, with speed T and the rate function J, uniformly for any initial measure in $\mathcal{M}_{\lambda_0,L}$ where $L > 1$ is any fixed number. Here $J : M_1(H) \to [0, +\infty]$ is the level-2 entropy of Donsker-Varadhan defined by (6.12) below.

More precisely we have:

i) J is a good rate function on $M_1(H)$ equipped with the topology τ of the convergence against bounded and borelian functions, i.e., $[J \le a]$ is τ-compact for every $a \in \mathbb{R}^+$.

ii) for all open set G in $M_1(H)$ with respect to the topology τ,

$$\liminf_{T\to\infty} \frac{1}{T} \log \inf_{\nu \in \mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(L_T \in G) \ge -\inf_G J \tag{6.4}$$

iii) for all closed set F in $M_1(H)$ with respect to the topology τ,

$$\limsup_{T\to\infty} \frac{1}{T} \log \sup_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(L_T \in F) \le -\inf_F J\ . \tag{6.5}$$

Furthermore we have

$$J(\nu) < +\infty \implies \nu \ll \mu,\ \nu(H_0^1) = 1 \text{ and } \int_{H_0^1} \|\nabla x\|_2^2 d\nu < +\infty \tag{6.6}$$

where μ is the unique invariant probability measure of (X_t).

The LDP w.r.t. the topology τ is much stronger than that w.r.t. the usual weak convergence topology as in Donsker-Varadhan [13]. Sometimes considered as a technical detail, the topology τ is crucial here : interesting consequences of this LDP can be deduced for many physical quantities of the system such as $\|x\|_{H^1} = \|\nabla x\|_2$, or more generally $\|x\|_{H^\alpha} := \|(-\Delta)^{\alpha/2}x\|_2$ for $0 \le \alpha \le 1$, which are not continuous on H. In fact, we establish

Corollary 6.1.2. *Let $\mathbb{B}$ a separable Banach space, and $f : H_0^1 \to \mathbb{B}$ a measurable function, bounded on the balls $\{x \text{ s.t } \|\nabla x\|_2 \le R\}$ for any $R > 0$, and satisfying*

$$\lim_{\|\nabla x\|_2\to\infty} \frac{\|f(x)\|_{\mathbb{B}}}{\|\nabla x\|_2^2} = 0. \tag{6.7}$$

Then, $\mathbb{P}_\nu\left(L_T(f) \in \cdot\right)$ satisfies the LDP on $\mathbb{B}$, with speed T and the rate function I_f given by

$$I_f(z) = \inf\{J(\nu);\ J(\nu) < +\infty, \int_{H_0^1} f(x)d\nu(x) = z\},\ \forall z \in \mathbb{B},$$

uniformly over initial distributions ν in $\mathcal{M}_{\lambda_0,L}$ for any fixed $L > 1$.

For instance, $f : H_0^1 \to \mathbb{B} := H^\alpha$ with $f(x) = x$ for any $\alpha \in [0,1)$ is allowed, so that the LDP in H^α holds for $\mathbb{P}_\nu\left(1/T\int_0^T X_t dt \in \cdot\right)$. An other particular case of the above corollary is the following: for every $p \in (0,2)$,

$$\mathbb{P}_\nu\left(\frac{1}{T}\int_0^T \|\nabla X(t)\|_2^p dt \in \cdot\right)$$

satisfies the LDP on $\mathbb{R}$ with speed T and the rate function I defined by

$$I(z) = \inf\left\{J(\nu); J(\nu) < +\infty, \int_H \|\nabla x\|_2^p d\nu(x) = z\right\},\ \forall z \in \mathbb{R} \tag{6.8}$$

uniformly over initial distributions ν in $\mathcal{M}_{\lambda_0,L}$ (for any $L > 1$).

Finally, we introduce $(e_k)_k$ the complete orthonormal system in $L^2(0,1)$ which diagonalizes Δ on its domain, and by $-\lambda_k$ the corresponding eigenvalues. We have

$$e_k(x) = \sqrt{\frac{2}{\pi}} \sin k\pi x, \quad \lambda_k = \pi^2 k^2, \quad k \in \mathbb{N}^* = \{1, 2, ...\}.$$

Remarks 6.1.3.

(i) Let us see the meaning of our assumptions: $tr(Q) < +\infty$ and (6.2). Assume that $Ge_k = \sigma_k e_k$ for every $k \geq 1$. Then

$$GW(t) = \sum_{k=1}^{\infty} \sigma_k \beta_k(t) e_k \tag{6.9}$$

where $(\beta_k)_{k \in \mathbb{N}^*}$ is a family of independent real valued standard Brownian motions. Then $tr(Q) < +\infty$ and condition (6.2) is satisfied if

$$\frac{c}{k} \leq |\sigma_k| \leq \frac{C}{k^{\frac{1}{2}+\varepsilon}}$$

for two positive constants c and C and some small $\varepsilon > 0$.

A more general example of noise for which our assumptions hold is

$$G := (-\Delta)^{-\beta} B, \ \frac{1}{4} < \beta < \frac{1}{2},$$

where B is any linear bounded and invertible operator on H. Indeed $tr(GG^*) \leq \|B\|^2_{H \to H} tr(\Delta^{-2\beta}) < +\infty$ for $2\beta > 1/2$. Since $Im(G) = Im(\Delta^{-\beta})$ and by the polar decomposition, $Im(G) \subset Im(\sqrt{GG^*})$, the condition (6.2) is then verified with $\delta = 2\beta$.

(ii) Our approach here is well adapted to the case of a multiplicative (or correlated) random forcing term, that is, the noise $GW(t)$ can be replaced by

$$g\left(X(t,\xi)\right) GW(t)$$

where $g : H \to [\alpha, \beta]$ is Lipschitz continuous, $0 < \alpha < \beta < \infty$, G satisfies (6.2) and $tr(GG^*) < +\infty$. Indeed, following [6], the strong Feller property and the topological irreducibility hold. All estimates necessary for the LDP in Theorem 6.1.1 still hold in the actual case, and then all previous results remain valid.

(iii) The class (6.3) of allowed initial distributions for the uniform LDP is sufficiently rich. For example, choosing L large enough, it includes all the Dirac probability measures δ_x with x in any ball of H.

(iv) Our LDP is more precise than the exponential convergence of P_t to the invariant measure μ established in [16]. Indeed the LDP furnishes the exact rate of the exponential convergence in probability of the empirical measures L_T to μ. Moreover by Theorem 6.4 in [24], under the strong Feller and topological irreducibility assumption for (P_t), the LDP in Theorem 6.1.1 is equivalent to saying that the essential spectral radius in some weighted functions spaces $b_u\mathcal{B}$ is zero.

(v) The assumption (6.2) plays a crucial role for Theorem 6.1.1: if the noise acts only on a finite number of modes (i.e., $\sigma_k = 0$ for all $k > N$ in (6.9)) as in Kolmogorov's turbulence theory, we believe that the LDP w.r.t. the τ-topology is false. It is a challenging open question for establishing the LDP of L_T w.r.t. the weak convergence topology in the last degenerate noise case.

(vi) For the 2D-stochastic Navier-Stokes equation, we can prove, under suitable conditions, a LDP on some $D(A^\alpha)$, for $\frac{1}{4} < \alpha < \frac{1}{2}$. Here A is also the Laplacian, but regarded as an operator on the subspace of the L^2-vector fields with free divergence. That will be carried out in a future work.

This paper is organized as follows. In Section 2, we recall known results on existence and uniqueness of solution, and existence of an invariant probability measure for equation (6.1). In section 3 we give some general facts about large deviations for strong Feller and irreducible Markov processes and we obtain the uniform lower bound (6.4). Then we prove the convergence of the Galerkin approximations for the considered equation in section 4. The exponential tightness is investigated in section 5, and the uniform upper bound (6.5) for the strong τ-topology in section 6. Finally, the extension to non bounded functionals on H is discussed in the last section 7.

6.2 Solutions of the equation and their properties

Let us specify what we understand by solution. Generally, we are concerned with two ways of giving a rigorous meaning to solutions of stochastic differential equations in infinite dimensional spaces, that is, the variational one [20], [19] and the semigroup one [8]. Correspondingly, as in the case of deterministic evolution equations, we have two notions of strong, and "mild" solution. In most situations, one finds that the concept of strong solution is too limited to include important examples. The weaker concept of mild solution seems to be more appropriate. In the sequel, we are working with this concept, that we define more precisely now.

We denote by $S(t)$ the semigroup generated by Δ on $L^2(0,1)$, or from a formal point of view, $S(t) = e^{t\Delta}$.

Definition 6.2.1. *We say that $X \in C\left([0,T], L^2(0,1)\right)$ is a "mild" solution of problem (6.1) if $X(t)$ is adapted to $\mathcal{F}_t$, the σ-algebra of the cylindrical Wiener process* *until time t and for arbitrary $0 \leq t$, we have*

$$X(t) = S(t)x_0 + \int_0^t S(t-s)\frac{1}{2}D_\xi X^2(s)ds + \int_0^t S(t-s)GdW(s) \tag{6.10}$$

for any $x_0 \in L^2(0,1)$, $\mathbb{P}$ almost surely.

Note that all the terms in (6.10) take sense since the mapping

$$F : u \in C\left([0,T], L^1(0,1)\right) \to \int_0^t S(t-s)\frac{1}{2}D_\xi u(s)ds \in C\left([0,T], L^2(0,1)\right)$$

is well defined (see [9] p260) and the stochastic convolution $W_\Delta := \int_0^t S(t-s)GdW(s)$ also (see (6.22) below). Da Prato, Debussche, Temam established in [7] for the first time existence and uniqueness for a stochastic Burgers equation cylindrically perturbed, that is when G is the identity operator. The method they used to obtain local existence in time of a solution consists in considering a fixed path of the noise, to get into a deterministic setting and use a fixed point argument. Then the time of explosion is shown to be infinite, by means of a priori bounds on the solution. The same proof gives in our setting:

Theorem 6.2.2. *Stochastic Burgers equation (6.1) admits a unique mild solution and for all $T > 0$,*

$$X \in C\left([0,T], L^2(0,1)\right) \cap L^2\left([0,T], C[0,1]\right).$$

The solution satisfies Markov and strong Markov properties (see [8]). We can also consider the transition semigroup associated to the dynamics given by

$$P_t\Phi(x) := \mathbb{E}\Phi\left(X(t,x)\right) = \mathbb{E}^x\Phi\left(X(t)\right), \ \forall \Phi \in b\mathcal{B}(H).$$

As in [7], this semigroup admits an invariant measure. Moreover, under our condition (6.2) on the noise, the following interesting properties hold.

Lemma 6.2.3. *(i) The transition semigroup (P_t) corresponding to the forced Burgers equation (6.1) satisfies the strong Feller property. That is, for any bounded borelian function Φ on H and any $t > 0$, the function $P_t\Phi(\cdot)$ is continuous on H.*

(ii) For every $t > 0$, $P_t(x, O) > 0$ for all $x \in H$ and all non-empty open subset O of H. Hence, (P_t) is also topologically irreducible.

(iii) In particular, the transition semigroup (P_t), corresponding to the forced Burgers equation (6.1) admits a unique invariant measure μ, which charges all non-empty open subsets of H.

Part (i) is well known when the cylindrical noise is considered (see [9]). In our case of a finite trace class noise, the non-degeneracy condition (6.2) is essential. More precisely, $\delta < 1$ allows to obtain a bound on the derivative of the semigroup by using the Bismut-Elworthy formula as in [4] or [15]. The condition $\delta > \frac{1}{2}$ is borrowed from the finite trace assumption, crucial in the application of Itô's formula for the exponential tightness.

The point (ii) was proved by Goldys and Maslowski in [16] for our class of noise. We recall that (P_t) is topologically irreducible if, for all non-empty open set Γ in H, and all $x \in H$, we have $P_t(x, \Gamma) > 0$ for some $t > 0$.

According to the general theory [9], we obtain (iii) as first corollary, sometimes called Doob's theorem, of the two preceding points together with the existence of invariant measure. In fact this result gives also the convergence of the transition probabilities to the invariant measure.

Our aim is to complete the study of the equation (6.1) by giving information on the rare events and the exact rate of exponential convergence by means of a large deviation principle, one of the strongest ergodic behaviors of Markov processes.

6.3 General results about large deviations

In this section, we introduce some necessary notations and definitions and give general results (essentially following [22]) on large deviations for Markov processes.

6.3.1 Notations and entropy of Donsker-Varadhan

We first compare the "topological irreducibility" defined above (often called irreducibility in the literature on SPDE) with the probabilistic irreducibility for a Markov process which is the more general assumption under which the large deviations result we use (as Lemma 6.3.2 below) holds true (see [18, 22] for details).

Let ν be a probability measure on H; a transition kernel operator P on H is said ν-irreducible (resp. ν-essentially irreducible) if for all A in H such that $\nu(A) > 0$, and for all x in H (resp. for ν almost all x in H), we can find $n \in \mathbb{N}$ such that $P^n(x, A) > 0$. When ν charges all non-empty open subsets of H, the ν-irreducibility implies the topological irreducibility. But for the strong Feller P, the topological irreducibility implies the ν-irreducibility for all ν such that $\nu \ll \nu P$ (see [22]).

Thus by Lemma 6.2.3, for the unique invariant measure μ of our model, P_t is μ-irreducible for every $t > 0$. In reality for our model, we have the much stronger property that all the probability measures in the family

$$\Big\{P_t(x, .),\ x \in H,\ t > 0\Big\}$$

are equivalent, and they are also equivalent to μ (see [9, p.41]).

Consider the $H := L^2(0,1)$-valued continuous Markov process

$$\left(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, (X_t(\omega))_{t\geq 0}, (\mathbb{P}_x)_{x\in H}\right)$$

whose semigroup of Markov transitions kernels is denoted by $(P_t(x,dy))_{t\geq 0}$, where

$\Omega = C(\mathbb{R}^+, H)$ is the space of continuous functions from $\mathbb{R}^+$ to H equipped with the compact convergence topology;

$\mathcal{F}_t = \sigma(X_s, 0 \leq s \leq t)$ for any $t \geq 0$ is the natural filtration;

$\mathcal{F} = \sigma(X_s, 0 \leq s)$ and $\mathbb{P}_x(X_0 = x) = 1$.

Hence, $\mathbb{P}_x$ is the law of the Markov process with initial state x in H. For any initial measure ν on H, let $\mathbb{P}_\nu(d\omega) := \int_H \mathbb{P}_x(d\omega)\nu(dx)$.

The empirical measure of level-3 (or process level) is

$$R_t := \frac{1}{t}\int_0^t \delta_{\theta_s X} ds$$

where $(\theta_s X)_t = X_{s+t}$ for all $t, s \geq 0$ are the shifts on Ω. Hence R_t is a random element of $M_1(\Omega)$, the space of probability measures on Ω.

The level-3 entropy functional of Donsker and Varadhan $H : M_1(\Omega) \to [0, +\infty]$ is defined by,

$$H(Q) := \begin{cases} \mathbb{E}^{\bar{Q}} h_{\mathcal{F}_1^0}\left(\bar{Q}_{\omega(-\infty,0]}; \mathbb{P}_{w(0)}\right), & \text{if } Q \in M_1^s(\Omega) \\ +\infty & \text{otherwise} \end{cases} \tag{6.11}$$

where

$M_1^s(\Omega)$ is the space of those elements in $M_1(\Omega)$ which are moreover stationary ;

$\bar{Q}$ is the unique stationary extension of $Q \in M_1^s(\Omega)$ to $\bar{\Omega} := C(\mathbb{R}, H)$;

$\mathcal{F}_t^s = \sigma(X(u); s \leq u \leq t)$ on $\bar{\Omega}$, $\forall s, t \in \mathbb{R}, s \leq t$;

$\bar{Q}_{\omega(-\infty,t]}$ is the regular conditional distribution of $\bar{Q}$ knowing $\mathcal{F}_t^{-\infty}$;

$h_{\mathcal{G}}(\nu,\mu)$ is the usual relative entropy or Kullback information of ν with respect to μ restricted on the σ-field $\mathcal{G}$, given by

$$h_{\mathcal{G}}(\nu,\mu) := \begin{cases} \int \frac{d\nu}{d\mu}|_{\mathcal{G}} \log\left(\frac{d\nu}{d\mu}|_{\mathcal{G}}\right) d\mu, & \text{if } \nu \ll \mu \text{ on } \mathcal{G} \\ +\infty & \text{otherwise.} \end{cases}$$

The level-2 entropy functional $J : M_1(H) \to [0,\infty]$ is defined by

$$J(\beta) = \inf\{H(Q)\ ;\ Q \in M_1^s(\Omega) \text{ and } Q_0 = \beta\},\ \forall \beta \in M_1(H), \tag{6.12}$$

where $Q_0(\cdot) = Q(X(0) \in \cdot)$ is the marginal law at time $t = 0$.

Lastly introduced in [22], we define the restriction of the Donsker Varadhan entropy to the μ component, by

$$H_\mu(Q) := \begin{cases} H(Q), & \text{if } Q_0 \ll \mu \\ +\infty & \text{otherwise} \end{cases}$$

and for the level-2 entropy functional

$$J_\mu(\beta) := \begin{cases} J(\beta), & \text{if } \beta \ll \mu \\ +\infty & \text{otherwise.} \end{cases}$$

For our model, let us first establish the

Lemma 6.3.1. *We have $J(\nu) < +\infty \implies \nu \ll \mu$. Moreover, $J = J_\mu$ on $M_1(H)$ and $[J = 0] = \{\mu\}$*

Proof. Consider ν such that $J(\nu) < \infty$. We recall the expression (6.11) of the Level-3 entropy. For $Q \in M_1^s(\Omega)$ such that $Q_0 = \nu$, $H(Q) < \infty$, and for every $t > 0$, noting that the entropy of marginal measure is not larger than the global entropy, we have by Jensen inequality,

$$\begin{aligned} H(Q) &= \mathbb{E}^{\bar{Q}} h_{\mathcal{F}_1}\left(\bar{Q}_{\omega(-\infty,0]}; \mathbb{P}_{w(0)}\right) \\ &= \frac{1}{t}\mathbb{E}^{\bar{Q}} h_{\mathcal{F}_t^0}\left(\bar{Q}_{\omega(-\infty,0]}; \mathbb{P}_{w(0)}\right) \\ &\geq \frac{1}{t} h_{\mathcal{F}_t^0}\left(Q; \mathbb{P}_\nu\right) \\ &\geq \frac{1}{t} h_{\sigma(w(t))}\left(Q; \mathbb{P}_\nu\right) \\ &\geq \frac{1}{t} h_{\mathcal{B}(H)}\left(\nu; \nu P_t\right). \end{aligned}$$

Taking infinimum over such Q, we get

$$J(\nu) \geq \frac{1}{t} h_{\mathcal{B}(H)}\left(\nu; \nu P_t\right). \tag{6.13}$$

So the Kullback information of ν with respect to νP_t is finite, which implies by definition that $\nu \ll \nu P_t$. Since all $P_t(x, dy), t > 0, x \in H$ are equivalent to μ ([9]), we have

$$\nu P_t(\cdot) = \int_H P(t, x, \cdot)\nu(dx) \ll \mu.$$

Thus $\nu \ll \nu P_t \ll \mu$, as desired.

By definition, we have $J \leq J_\mu$ and they are equal on

$$\{\nu \in M_1(H) \ \text{ such that } \ \nu \ll \mu\}.$$

Since any probability measure ν on H such that $J(\nu) < \infty$ is absolutely continuous with respect to μ, we have $J = J_\mu$ on $M_1(H)$.

At the end, if the probability measure β is such that $J(\beta) = 0$ then $\beta \ll \mu$ and $\beta = \beta P_t$ for every $t > 0$ by (6.13). By the uniqueness in Lemma 6.2.3, we have $\beta = \mu$ and the proof is finished. □

6.3.2 The lower bound

Let us first recall the definition of the projective limit τ_p of the strong τ-topology,

$$\tau_p := \sigma\left(M_1(\Omega), \cup_{t\geq 0} b\mathcal{F}_t^0\right)$$

where $b\mathcal{F}_t^0$ is the set of functions on Ω, that are bounded and measurable for $\mathcal{F}_t^0$.

The following level-3 lower bound of Large Deviations for τ_p was established by Wu (see [22, th. B.1]) under more general conditions.

Lemma 6.3.2. ([22]) *For any open set O in $(M_1(\Omega), \tau_p)$,*

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}_x(R_t \in O) \geq -\inf_O H_\mu, \quad \mu - a.e \text{ initial state } x \in H.$$

Recall that $H_\mu = H$ by Lemma 6.3.1. Our goal here, is to prove the

Proposition 6.3.3. *If J is a good rate function on $(M_1(H), \tau)$ and the uniform upper bound (6.5) is satisfied, then the level-3 uniform lower bound holds true: for any measurable open subset O in $(M_1(\Omega), \tau_p)$,*

$$\liminf_{t\to\infty} \frac{1}{t} \log \inf_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(R_t \in O) \geq -\inf_O H.$$

In particular, the desired Level-2 lower bound (6.4) holds (by the contraction principle).

Proof. For any $Q \in O$ fixed, we can take a τ_p neighborhood of Q in $M_1(\Omega)$ of form

$$N(Q,\delta) := \{Q' \in M_1(\Omega) \text{ such that } \left|\int F_i dQ' - \int F_i dQ\right| < \delta, \ \ \forall i = 1, \cdots, d\}$$

contained in O, where $\delta > 0$, $1 \leq d \in \mathbb{N}$ and $F_i \in b\mathcal{F}_n^0$ for some $n \in \mathbb{N}$. It is sufficient to establish that for every Q in O such that $H(Q) < \infty$

$$\liminf_{t\to\infty} \frac{1}{t} \log \inf_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(R_t \in N(Q,\delta)) \geq -H_\mu(Q). \tag{6.14}$$

But by Egorov's lemma, Lemma 6.3.2 implies the existence of a borelian subset K in H with $\mu(K) > 0$ such that for any $\varepsilon > 0$

$$\inf_{x\in K} \frac{1}{t}\log \mathbb{P}_x\left(R_t \in N\left(Q,\frac{\delta}{2}\right)\right) \geq -H_\mu(Q) - \varepsilon \tag{6.15}$$

for all t large enough. Let us fix $a > 0$. For any $0 \leq b \leq a$, we have

$$\left|\int F_i d(R_t\circ\theta_b - R_t)\right| \leq \frac{2(a+1)}{t}\|F_i\|_\infty$$

and then for all $0 \leq b \leq a$ and for all t large enough (depending on a and δ),

$$\begin{aligned}\mathbb{P}_\nu(R_t \in N(Q,\delta)) \;&\geq\; \mathbb{P}_\nu\left(X_b \in K \;\;;\;\; R_t\circ\theta_b \in N\left(Q,\frac{\delta}{2}\right)\right)\\ &\geq\; \mathbb{P}_\nu(X_b \in K)\;\;\inf_{x\in K}\mathbb{P}_x\left(R_t \in N\left(Q,\frac{\delta}{2}\right)\right).\end{aligned}$$

Integrating for $0 \leq b \leq a$, and dividing by a yields

$$\mathbb{P}_\nu(R_t \in N(Q,\delta)) \geq \mathbb{E}^\nu L_a(K)\;\;\inf_{x\in K}\mathbb{P}_x\left(R_t \in N\left(Q,\frac{\delta}{2}\right)\right). \tag{6.16}$$

Hence, for proving (6.14), by (6.15) and (6.16), it is enough to establish that for any borelian subset K with $\mu(K) > 0$, we can find $a > 0$ such that

$$\inf_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{E}^\nu L_a(K) > 0.$$

Notice that

$$\mathbb{E}^\nu L_a(K) \geq \frac{\mu(K)}{2}\left(1 - \mathbb{P}_\nu\left(|L_a(K) - \mu(K)| \geq \frac{\mu(K)}{2}\right)\right)$$

and by the assumed level 2 upper bound,

$$\limsup_{a\to+\infty}\frac{1}{a}\log\sup_{\nu\in\mathcal{M}_{\lambda_0,L}}\mathbb{P}_\nu\left(|L_a(K)-\mu(K)| \geq \frac{\mu(K)}{2}\right) \leq -\inf_F J(\nu)$$

where

$$F = \left\{\beta \in M_1(H) \;:\; |\beta(K) - \mu(K)| \geq \frac{\mu(K)}{2}\right\}$$

is closed for the τ-topology. So once $\inf_F J > 0$, we shall obtain for all a large enough

$$\inf_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{E}^\nu L_a(K) \geq \frac{\mu(K)}{2}\left(1 - \exp\left(-a\frac{\inf_F J}{2}\right)\right) > 0.$$

It remains to prove that $\inf_F J > 0$. To this end we may assume that $\inf_F J < +\infty$. In that case, since J is a good rate function (our condition), $\inf_F J$ is attained by some $\beta_0 \in F$. But $J(\beta) = 0 \Leftrightarrow \beta = \mu$ (Lemma 6.3.1) and $\mu \notin F$, so $\inf_F J = J(\beta_0) > 0$ as desired. $\square$

6.3.3 Cramer functionals and weak upper bound

Let us introduce the uniform upper Cramer functional over a non-empty family of initial measures $\mathcal{A}$ in $M_1(H)$,

$$\Lambda(V|\mathcal{A}) := \limsup_{t\to\infty} \frac{1}{t} \log \sup_{\nu\in\mathcal{A}} \mathbb{E}^{\nu} \exp\left(tL_t(V)\right)$$

and several other Cramer functionals,

$$\begin{aligned}
\Lambda(V|x) &:= \Lambda(V|\{\delta_x\}) = \limsup_{t\to\infty} \frac{1}{t} \log \mathbb{E}^{x} \exp\left(tL_t(V)\right) \\
\Lambda^{\infty}(V) &:= \Lambda(V|\{\delta_x\ ;\ x\in H\}) = \limsup_{t\to\infty} \frac{1}{t} \log \sup_{x\in H} \mathbb{E}^{x} \exp\left(tL_t(V)\right) \\
\Lambda^{0}(V) &:= \sup_{x\in H} \Lambda(V|x) = \sup_{x\in H} \limsup_{t\to\infty} \frac{1}{t} \log \mathbb{E}^{x} \exp\left(tL_t(V)\right) \qquad (6.17)
\end{aligned}$$

where V is a bounded and borelian function on H.

The functionals $\Lambda^0(V)$ and $\Lambda^\infty(V)$ are respectively the pointwise and uniform Cramer functionals introduced already in [12]. For $\Lambda : b\mathcal{B}(H) \to \mathbb{R}$ any one of the above functionals, define its Legendre transformation:

$$\begin{aligned}
\Lambda_w^*(\nu) &= \sup_{V\in C_b(H)} \left(\int_H V d\nu - \Lambda(V) \right), \quad \forall \nu \in M_b(H) \\
\Lambda^*(\nu) &= \sup_{V\in b\mathcal{B}(H)} \left(\int_H V d\nu - \Lambda(V) \right), \quad \forall \nu \in M_b(H) \qquad (6.18)
\end{aligned}$$

where $M_b(H)$ is the space of all signed σ-additive measures of bounded variation on $(H, \mathcal{B})$.

Remark that $\{\delta_x\}_{x\in H} \subset \cup_{L>0} \mathcal{M}_{\lambda_0,L}$, we have for any bounded and measurable function V,

$$\Lambda^0(V) \le \sup_{L>0} \Lambda(V|\mathcal{M}_{\lambda_0,L}) \le \Lambda^\infty(V).$$

Since (P_t) is Feller, we have by [22, Proposition B.13]

$$\left(\Lambda^0\right)^*(\nu) = \left(\Lambda^0\right)_w^*(\nu) = \left(\Lambda^\infty\right)^*(\nu) = \left(\Lambda^\infty\right)_w^*(\nu) = J(\nu),\ \forall \nu \in M_1(H)$$

which implies the l.s.c for J and the fact that

$$\begin{aligned}
&\sup_{V\in b\mathcal{B}(H)} \left(\int_H V d\nu - \sup_{L>0} \Lambda(V|\mathcal{M}_{\lambda_0,L}) \right) \\
&= \sup_{V\in C_b(H)} \left(\int_H V d\nu - \sup_{L>0} \Lambda(V|\mathcal{M}_{\lambda_0,L}) \right) = J(\nu),\ \forall \nu \in M_1(H).
\end{aligned}$$

So by Gärtner-Ellis theorem (see [10]), we have always the following general weak* upper bound

Lemma 6.3.4. *Let $M_1(H)$ be equipped with the weak convergence topology. For any compact subset K in $M_1(H)$ w.r.t the weak convergence topology, and for any $\varepsilon > 0$, there is a neighborhood $N(K, \varepsilon)$ of K in $M_1(H)$ such that*

$$\limsup_{t\to\infty} \frac{1}{t} \log \sup_{\nu \in \mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(L_t \in N(K,\varepsilon))$$
$$\leq \begin{cases} -\inf_{\nu\in K} J(\nu) + \varepsilon & \textit{if} \quad \inf_{\nu\in K} J(\nu) < \infty \\ -\frac{1}{\varepsilon} & \textit{otherwise.} \end{cases}$$

Now to obtain the upper bound in Theorem 6.1.1 w.r.t. the weak convergence topology, we need to prove the exponential tightness of L_t.

6.4 Convergence of a Galerkin method

Let us introduce the approximation system associated with equation (6.1) :

$$dX_n(t) = \left(\Delta X_n(t) + \frac{1}{2}\Pi_n D_\xi (X_n)^2(t)\right) dt + G_n dW(t) \; ; \; X_n(0) = \Pi_n x \tag{6.19}$$

where Π_n is the orthogonal projection on H_n, the finite dimensional space spanned by the first n eigenvectors $(e_1, \ldots, e_n)$, and $G_n := \Pi_n G$.

The convergence of a similar approximation but with a non linearity truncated by the function

$$f_n(x) = \frac{nx^2}{n + x^2}$$

was investigated by Da Prato and Debussche [5]. The aim of this section is to establish some *a priori* estimates on X_n, and the convergence of the approximation method (6.19).

Theorem 6.4.1. *The solutions X_n of (6.19) converge to the solution X of (6.1) in $C([0,T]; H)$ and in $L^2([0,T]; H_0^1)$ almost surely.*

6.4.1 A priori estimate for the finite dimensional approximations

From now on, we denote by $\langle \cdot, \cdot \rangle$ the inner product in H. Let us apply Itô's formula to the finite dimensional diffusion X_n. Since $X_n(t) \in H_n$, remark that

$$\langle X_n(t), \Pi_n D_\xi X_n^2(t)\rangle \;=\; \langle \Pi_n X_n(t), D_\xi X_n^2(t)\rangle = \langle X_n(t), D_\xi X_n^2(t)\rangle$$

$$= \int_0^1 X_n(t,\xi) D_\xi X_n^2(t,\xi) d\xi = \left[\frac{X_n^3(t,\xi)}{3}\right]_{\xi=0}^{\xi=1} = 0$$

because of no-slip boundary conditions. So, we obtain:

$$\begin{aligned} d\|X_n(t)\|_2^2 &= 2\langle X_n(t), dX_n(t)\rangle + tr(Q_n)dt \\ &= \left[-2\|\nabla X_n(t)\|_2^2 + tr(Q_n)\right] dt + 2\langle X_n(t), G_n dW(t)\rangle. \end{aligned}$$

In the same spirit, denoting by $d[Y,Y]_t$ the quadratic variation process of a semi-martingale Y, we can also compute with the Itô formula

$$\begin{aligned} de^{\lambda_0\|X_n(t)\|_2^2} &= e^{\lambda_0\|X_n(t)\|_2^2}\left[\lambda_0 d\|X_n(t)\|_2^2 + \frac{\lambda_0^2}{2} d[\|X_n\|_2^2, \|X_n\|_2^2]_t\right] \\ &= e^{\lambda_0\|X_n(t)\|_2^2}\left[-2\lambda_0\|\nabla X_n(t)\|_2^2 + \lambda_0 tr(Q_n) + 2\lambda_0^2\|G_n^* X_n(t)\|_2^2\right] dt \\ &\quad +2\lambda_0 e^{\lambda_0\|X_n(t)\|_2^2}\langle X_n(t), G_n dW(t)\rangle. \end{aligned}$$

For any smooth function f on $H_n = \Pi_n(L^2)$, we define $g := \mathcal{L}_n f$ if

$$f(X_n(t)) - f(X_n(0)) - \int_0^t g(X_n(s))ds$$

is a local martingale. The following lemma, being a consequence of Itô's formula, is well known to probabilists and it is crucial.

Lemma 6.4.2. ([21]) *If f is smooth on H_n, and $f \geq 1$, then*

$$M_t := e^{-\int_0^t \frac{\mathcal{L}_n f}{f}(X_n(s))ds} f(X_n(t))$$

is a local martingale.

In view of the above definition we have for $f(x) = e^{\lambda_0\|x\|_2^2}$ $(x \in H_n)$,

$$\mathcal{L}_n f(x) := f(x)\left[-2\lambda_0\|\nabla x\|_2^2 + \lambda_0 tr(Q_n) + 2\lambda_0^2\|G_n^* x\|_2^2\right]$$

and

$$\begin{aligned} -\frac{\mathcal{L}_n f(x)}{f(x)} &= 2\lambda_0\|\nabla x\|_2^2 - \lambda_0 tr(Q_n) - 2\lambda_0^2\|G_n^* x\|_2^2 \\ &\geq 2\lambda_0\|\nabla x\|_2^2 - \lambda_0 tr(Q) - 2\lambda_0^2\|Q\|\ \|x\|_2^2. \end{aligned}$$

Moreover, by the Poincaré inequality

$$\|x\|_2 \leq \frac{\|\nabla x\|_2}{\pi},\ \forall x \in H$$

we obtain for $0 < \lambda_0 \leq \frac{\pi^2}{2\|Q\|}$, since $1 - \frac{\lambda_0\|Q\|}{\pi^2} \geq \frac{1}{2}$

$$\begin{aligned} -\frac{\mathcal{L}_n f(x)}{f(x)} &\geq 2\lambda_0 \left(\|\nabla x\|_2^2 \left(1 - \frac{\lambda_0\|Q\|}{\pi^2}\right) - \frac{tr(Q)}{2} \right) \\ &\geq \lambda_0\|\nabla x\|_2^2 - \lambda_0 tr(Q). \end{aligned}$$

So we conclude by Lemma 6.4.2 that

$$N_t^n := \exp\left(\lambda_0 \int_0^t \|\nabla X_n(s)\|_2^2 ds - \lambda_0 tr(Q)t\right) e^{\lambda_0\|X_n(t)\|_2^2} \tag{6.20}$$

is a supermartingale. This proves the following crucial exponential estimate:

Lemma 6.4.3. *Let $0 < \lambda_0 < \frac{\pi^2}{2\|Q\|}$. For any x in H, we have*

$$\mathbb{E}^x \exp\left(\lambda_0 \int_0^t \|\nabla X_n(s)\|_2^2 ds\right) e^{\lambda_0\|X_n(t)\|_2^2} \leq e^{\lambda_0 tr(Q)t} e^{\lambda_0\|x\|_2^2}, \ \forall t > 0. \tag{6.21}$$

In particular, we have

$$\sup_{n\in\mathbb{N}} \mathbb{E}^x e^{\lambda_0 \int_0^t \|\nabla X_n(s)\|_2^2 ds} < \infty$$

so $(X_n)_n$ is uniformly bounded in $L^2(\Omega \times [0,T], H_0^1)$.

This kind of properties was also investigated by Da Prato and Debussche [4] for proving some estimates on derivatives of the transition semigroup.

6.4.2 Proof of theorem 6.4.1

Let us introduce the stochastic convolution, or Ornstein-Uhlenbeck process

$$W_\Delta(t) = \int_0^t S(t-s)GdW(s)$$

which is the mild solution of the linear equation with additive noise

$$dW_\Delta(t) = \Delta W_\Delta(t)dt + GdW(t) \ ; \ W_\Delta(0) = 0. \tag{6.22}$$

Since $Q = GG^*$ has finite trace, it is known (see [8, p.99 and p.148]), that the stochastic integral W_Δ is the limit in $L^2(\Omega, C([0,T];H))$ and in $L^2(\Omega, L^2([0,T];H_0^1))$ of its finite dimensional approximation defined by

$$W_\Delta^n(t) = \int_0^t S(t-s)G_n dW(s) = \Pi_n W_\Delta(t).$$

Notice that W_Δ^n is the mild solution of the finite dimensional linear equation with additive noise

$$dW_\Delta^n(t) = \Delta W_\Delta^n(t)dt + G_n dW(t)\ ;\ W_\Delta(0) = 0. \tag{6.23}$$

Let us prove that the convergences above hold in fact a.s. in $C([0,T];H)$ and $L^2([0,T];H_0^1)$. Indeed the a.s. convergence of W_Δ^n to W_Δ in $L^2([0,T];H_0^1)$ is obvious. For the convergence in $C([0,T],H)$, since for a.e. ω, $t \mapsto W_\Delta(t,\omega)$ is continuous from $[0,T]$ to H, then $K := \{W_\Delta(t,\omega); t \in [0,T]\}$ is compact in H. Notice that if $h \in H$, $\Pi_n h \to h$ in H and that the mappings $h \mapsto \Pi_n h,\ n \geq 1$ are equi-continuous on H for $\|\Pi_n\|_{H\to H} = 1$. So the above pointwise convergence is uniform over the compact subset K by Arzela-Ascoli's theorem : as $n \to \infty$,

$$\sup_{t\in[0,T]} \|W_\Delta(t,\omega) - \Pi_n W_\Delta(t,\omega)\|_H \to 0.$$

Our proof below, as in [9], will be completely deterministic. Fix any $\omega \in \Omega$ such that $W_\Delta^n(\omega) \to W_\Delta(\omega)$ both in $C([0,T],H)$ and $L^2([0,T],H_0^1)$ and we shall remove "ω" in the proof below.

Let us define

$$\begin{aligned} y &:= X - W_\Delta = S(t)x + \frac{1}{2}\int_0^t S(t-s)D_\xi(X)^2(s)ds \\ y_n &:= X_n - W_\Delta^n = S(t)\Pi_n x + \frac{1}{2}\int_0^t S(t-s)\Pi_n D_\xi(X_n)^2(s)ds \end{aligned}$$

and

$$z_n := y - y_n = X - X_n - (W_\Delta - W_\Delta^n).$$

Recall that X is bounded in $L^\infty([0,T],H)$ and also in $L^2([0,T],H_0^1)$ almost surely. Indeed we have the following a-priori estimates (see [9, p264]) :

$$\|y(t)\|_2^2 \leq e^{8\int_0^t \|W_\Delta(s)\|_\infty^2 ds}\|x\|_2^2 + 2\int_0^t e^{8\int_r^t \|W_\Delta(s)\|_\infty^2 ds}\|W_\Delta(r)\|_\infty^4 dr$$

and

$$\int_0^T \|D_\xi y(t)\|_2^2 dt \leq 8\int_0^T \|W_\Delta(t)\|_\infty^2 \|y(t)\|_2^4 dt + \int_0^T \|W_\Delta(t)\|_\infty^4 dt$$

and the fact that $H_0^1 \subset C[0,1]$ is a compact continuous embedding. The same proof as in [9, p264] yields the same estimates for y_n with W_Δ replaced by W_Δ^n, so the sequence $(X_n)_n$ is bounded in $L^\infty([0,T],H)$ and $L^2([0,T],H_0^1)$ almost surely (see also [5]). We can assume without loss of generality that the preceding bounds hold for our "ω".

It remains to show the convergence of z_n to 0 in the desired spaces. Notice that z_n is solution of

$$\frac{dz_n}{dt} = \Delta z_n + \frac{1}{2}D_\xi X^2 - \frac{1}{2}\Pi_n D_\xi X_n^2$$

from which we can deduce the a priori estimate:

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\|z_n\|_2^2 + \|D_\xi z_n\|_2^2 &= \langle\frac{1}{2}D_\xi(X)^2 - \frac{1}{2}\Pi_n D_\xi(X_n)^2, z_n\rangle \\
&= \langle\frac{1}{2}D_\xi(X^2 - X_n^2), z_n\rangle + (\frac{1}{2}(I-\Pi_n)D_\xi(X^2), z_n) \\
&:= I_1 + I_2.
\end{aligned}$$

Noting that $X - X_n = z_n + (I - \Pi_n)W_\Delta$, we have

$$\begin{aligned}
I_1 &= -\frac{1}{2}\langle(X_n + X)(X - X_n), D_\xi z_n\rangle \\
&= -\frac{1}{2}\langle(X_n + X)z_n, D_\xi z_n\rangle - \frac{1}{2}\langle(X_n + X)(I - \Pi_n)W_\Delta, D_\xi z_n\rangle \\
&:= I_{11} + I_{12}.
\end{aligned}$$

We can bound I_{11} as follows

$$\begin{aligned}
|I_{11}| &\leq \frac{1}{2}\|X_n + X\|_\infty\|z_n\|_2\|D_\xi z_n\|_2 \\
&\leq \frac{1}{4}\|D_\xi z_n\|_2^2 + \|X_n + X\|_\infty^2\|z_n\|_2^2
\end{aligned}$$

and for the second

$$\begin{aligned}
|I_{12}| &\leq \frac{1}{2}\|X_n + X\|_\infty\|(I - \Pi_n)W_\Delta\|_2\|D_\xi z_n\|_2 \\
&\leq \frac{1}{4}\|D_\xi z_n\|_2^2 + \|X_n + X\|_\infty^2\|(I - \Pi_n)W_\Delta\|_2^2.
\end{aligned}$$

Similarly, for the remaining term, we have

$$|I_2| \leq \|(I - \Pi_n)D_\xi(X)\|_2^2 + \|X\|_\infty^2\|z_n\|_2^2.$$

Hence we obtain the inequality

$$\begin{aligned}
\frac{d}{dt}\|z_n\|_2^2 + \|D_\xi z_n\|_2^2 \leq\ & 2\left(\|X\|_\infty^2 + \|X_n + X\|_\infty^2\right)\|z_n\|_2^2 + 2\|(I - \Pi_n)D_\xi(X)\|_2^2 \\
&+2\|X_n + X\|_\infty^2\|(I - \Pi_n)W_\Delta\|_2^2. \qquad (6.24)
\end{aligned}$$

By Gronwall's inequality we get

$$\|z_n(t)\|_2^2 \leq \exp\left(\int_0^t 2\|X(s)\|_\infty^2 + 2\|X_n(s) + X(s)\|_\infty^2 ds\right)\|(I - \Pi_n)x\|_2^2$$

$$+2\int_0^t \exp\left(\int_s^t 2\|X(r)\|_\infty^2 + 2\|X_n(r)+X(r)\|_\infty^2 dr\right)\|X_n(s)+X(s)\|_\infty^2\|(I-\Pi_n)W_\Delta(s)\|_2^2 ds$$
$$+2\int_0^t \exp\left(\int_s^t 2\|X(r)\|_\infty^2 + 2\|X_n(r)+X(r)\|_\infty^2 dr\right)\|(I-\Pi_n)D_\xi X(s)\|_2^2 ds.$$

In the sequel we denote the norm in the corresponding spaces respectively by

$|u|^2_{L^2(0,T,H)} := \int_0^T \|u(t)\|_2^2 dt$

$|u|^2_{L^2(0,T,H_0^1)} := \int_0^T \|\nabla u(t)\|_2^2 dt = \int_0^T \|D_\xi u(t)\|_2^2 dt$

$|u|_{C(0,T,H)} := \sup_{0\le t\le T}\|u(t)\|_2$

Taking the supremum in t, and using again the compact continuous embedding $H_0^1(0,1)\subset C[0,1]$, so that $\|x\|_\infty \le C\|\nabla x\|_2$ for some constant $C>0$, we obtain

$$\begin{aligned}\sup_{0\le t\le T}\|z_n(t)\|_2^2 &\le e^{\left(4C^2|X_n|^2_{L^2(0,T,H_0^1)}+6C^2|X|^2_{L^2(0,T,H_0^1)}\right)}\Big((1)+(2)+(3)\Big)\\ &\le e^{M_1}\Big((1)+(2)+(3)\Big) \qquad (6.25)\end{aligned}$$

for some number $M_1>0$, where

$$\begin{aligned}(1) &= \|(I-\Pi_n)x\|_2^2\\ (2) &= 4C^2\left(|X_n|^2_{L^2(0,T,H_0^1)}+|X|^2_{L^2(0,T,H_0^1)}\right)|(I-\Pi_n)W_\Delta|^2_{C(0,T,H)}\\ &\le M_2|(I-\Pi_n)W_\Delta|^2_{C(0,T,H)}\\ (3) &= 2\int_0^T\|(I-\Pi_n)D_\xi X(s)\|_2^2 ds.\end{aligned}$$

for some constant $M_2>0$.

Now, $(1)\to 0$ is clear, $(2)\to 0$ is asssumed for our "ω", and $(3)\to 0$ by dominate convergence. Consequently, $z_n\to 0$ in $C([0,T],H)$.

Finally, let us integrate (6.24) for t. It gives

$$\begin{aligned}|z_n|^2_{L^2([0,T],H_0^1)} &\le \left(4C^2|X_n|^2_{L^2([0,T],H_0^1)}+6C^2|X|^2_{L^2([0,T],H_0^1)}\right)\sup_{0\le t\le T}\|z_n(t)\|_2^2\\ &\quad +2|(I-\Pi_n)D_\xi(X)|^2_{L^2([0,T],H)}+\|(I-\Pi_n)x\|_2^2\\ &\quad +4C^2\left(|X_n|^2_{L^2([0,T],H_0^1)}+|X|^2_{L^2([0,T],H_0^1)}\right)|(I-\Pi_n)W_\Delta|^2_{C([0,T],H)}\\ &\le M_1\sup_{0\le t\le T}\|z_n(t)\|_2^2+2|(I-\Pi_n)D_\xi(X)|^2_{L^2([0,T],H)}\\ &\quad +\|(I-\Pi_n)x\|_2^2+M_2|(I-\Pi_n)W_\Delta|^2_{C([0,T],H)}\end{aligned}$$

which yields $z_n\to 0$ in $L^2(0,T,H_0^1)$ and the proof is finished. □

6.5 Uniform upper bound for the weak convergence topology: the exponential tightness

In this section, $M_1(H)$ is equipped with $\sigma(M_1(H), C_b(H))$ the weak convergence topology , instead of τ. The aim is to prove the following

Proposition 6.5.1. *(a) For any $\varepsilon > 0$, there is some compact subset $K = K_\varepsilon$ in $M_1(H)$ in the weak convergence topology such that*

$$\limsup_{t\to\infty} \frac{1}{t} \log \sup_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(L_t \notin K) \le -\frac{1}{\varepsilon}.$$

(b) Consequently for any closed set F in $M_1(H)$ equipped with the weak convergence topology $\sigma(M_1(H), C_b(H))$,

$$\limsup_{t\to\infty} \frac{1}{t} \log \sup_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu(L_t \in F) \le -\inf_F J.$$

By the weak upper bound in Lemma 6.3.4 and according to the general theory of large deviations, the upper bound of large deviations in (b) follows from the uniform exponential tightness of the family of $\mathbb{P}_\nu(L_T \in \cdot)$ over $\nu \in \mathcal{M}_{\lambda_0,L}$ stated in part (a).

Before proving it, let us notice the following consequence of our study in §4.

Lemma 6.5.2. *For any fixed $0 < \lambda_0 < \frac{\pi^2}{2\|Q\|}$,*

$$N_t := \exp\left(\lambda_0 \int_0^t \|\nabla X(s)\|_2^2 ds - \lambda_0 tr(Q)t\right) e^{\lambda_0\|X(t)\|_2^2}$$

is a supermartingale. In particular we have

$$\mathbb{E}^x e^{\lambda_0 \int_0^t \|\nabla X(s)\|_2^2 ds} \le e^{\lambda_0 tr(Q)t} e^{\lambda_0\|x\|_2^2}, \quad \forall x \in H \tag{6.26}$$

and for any fixed $L > 1$, and any initial measure in the set $\mathcal{M}_{\lambda_0,L}$, the following estimate holds

$$\mathbb{E}^\nu e^{\lambda_0 \int_0^t \|\nabla X(s)\|_2^2 ds} \le e^{\lambda_0 tr(Q)t} L. \tag{6.27}$$

Proof. By the almost sure convergence in Theorem 6.4.1 and Fatou's Lemma, (N_t) is a supermartingale by passing to the limit for $n \to \infty$ in (6.20). The estimates (6.26) and (6.27) follow immediately. □

Proof of Proposition 6.5.1. As said above it is sufficient to prove the uniform exponential tightness of $(\mathbb{P}_\nu(L_t \in \cdot), t \to +\infty)$ over $\nu \in \mathcal{M}_{\lambda_0,L}$ in part (a).

Step 1. Define $\Phi : M_1(H) \to [0, +\infty]$ by

$$\Phi(\beta) = \lambda_0 \int_H \|\nabla x\|_2^2 \, d\beta(x), \text{ with } \|\nabla x\|_2 := +\infty \text{ for } \forall x \in H \backslash H_0^1$$

where λ_0 is a real number such that $0 < \lambda_0 < \frac{\pi^2}{2\|Q\|}$. We claim that this function admits compact level sets.

At first, $x \to \|\nabla x\|_2^2$ is lower semi continuous (l.s.c. in short) on H, as a non decreasing limit of continuous functions $x \to \|\nabla \Pi_n x\|_2^2$. Thus, Φ is l.s.c. on $M_1(H)$, and for any $a > 0$, the level set $[\Phi \le a]$ is closed in $M_1(H)$.

Now let us show that $[\Phi \le a]$ is tight (thus compact in $M_1(H)$ by Prokorov's criterion). Indeed, for any $\delta > 0$ consider

$$A_\delta = \left\{ x \in H_0^1 \ s.t. \ \|\nabla x\|_2 \le \sqrt{\frac{a}{\lambda_0 \delta}} \right\}.$$

It is compact in H by the compact embedding $H_0^1 \subset H$, and we have

$$\forall \beta \in [\Phi \le a], \ \beta(A_\delta^c) \le \int_{A_\delta^c} \frac{\lambda_0 \delta \|\nabla x\|_2^2}{a} \, d\beta(x) \le \delta \frac{\Phi(\beta)}{a} \le \delta.$$

Step 2. For any $\varepsilon > 0$, $K := [\Phi \le \lambda_0 tr(Q) + 1/\varepsilon]$ is a compact subset of $M_1(H)$ by Step 1. For any $\nu \in \mathcal{M}_{\lambda_0, L}$, we have by Chebychev's inequality and Lemma 6.5.2,

$$\begin{aligned}
\mathbb{P}_\nu(L_t \notin K) &\le \exp\left(- \left[\lambda_0 tr(Q) + \frac{1}{\varepsilon} \right] t \right) \mathbb{E}^\nu e^{t\Phi(L_t)} \\
&= \exp\left(- \left[\lambda_0 tr(Q) + \frac{1}{\varepsilon} \right] t \right) \mathbb{E}^\nu \exp\left(\lambda_0 \int_0^t \|\nabla X(s)\|_2^2 ds \right) \\
&\le e^{-t/\varepsilon} L,
\end{aligned}$$

the desired uniform exponential tightness. □

6.6 Uniform upper bound for the τ-topology

Now, we prove the desired Upper Bound (6.5) for the strong τ-topology. It is based on the following criterion of the so-called hyper-exponential recurrence [23, Theorem 2.1] established by Wu for strong Feller and topologically irreducible Markov processes.

Lemma 6.6.1. ([23]) *For a subset K in H, let us define $\tau_K := \inf\{t \ge 0 \ s.t \ X_t \in K\}$ and $\tau_K^{(1)} := \inf\{t \ge 1 \ s.t \ X_t \in K\}$. If for any $\lambda > 0$, there exists a compact subset K in H such that*

$$\sup_{\nu \in \mathcal{M}_{\lambda_0, L}} \mathbb{E}^\nu e^{\lambda \tau_K} < \infty \tag{6.28}$$

and

$$\sup_{x\in K}\mathbb{E}^x e^{\lambda\tau_K^{(1)}}<\infty \tag{6.29}$$

then $[J\leq a]$ is τ-compact for every $a\in\mathbb{R}^+$, and the Upper Bound (6.5) uniform on $\mathcal{M}_{\lambda_0,L}$ for the τ-topology holds true.

In this section we establish the estimates (6.28) and (6.29) for our model. For the compact subset K of H, we still consider

$$K:=\left\{x\in H_0^1\ s.t\ \|\nabla x\|_2\leq M\right\} \tag{6.30}$$

where the real number M will be fixed later. The definition of the occupation measure implies that for $n\geq 2$,

$$\mathbb{P}_\nu\left(\tau_K^{(1)}>n\right)\leq\mathbb{P}_\nu\left(L_n(K)\leq\frac{1}{n}\right)=\mathbb{P}_\nu\left(L_n(K^c)\geq 1-\frac{1}{n}\right).$$

With our choice for K, we remark that

$$L_n(K^c)\leq\frac{1}{M^2}L_n(\|\nabla x\|_2^2).$$

Hence for any fixed real $0<\lambda_0<\frac{\pi^2}{2\|Q\|}$, we have by Chebychev's inequality

$$\begin{aligned}\mathbb{P}_\nu\left(\tau_K^{(1)}>n\right) &\leq \mathbb{P}_\nu\left(L_n(\|\nabla x\|_2^2)>M^2\left(1-\frac{1}{n}\right)\right)\\ &\leq e^{-n\lambda_0M^2\left(1-\frac{1}{n}\right)}\mathbb{E}^\nu e^{\lambda_0\int_0^n\|\nabla X_s\|_2^2ds}.\end{aligned}$$

For any initial measure $\nu\in M_1(H)$, integrating (6.26) w.r.t ν, and using it in the above expression yields

$$\mathbb{P}_\nu\left(\tau_K^{(1)}>n\right)\ \leq\ \int_H e^{\lambda_0\|x\|_2^2}\nu(dx)\ \ e^{-n\lambda_0C},\ \forall n\geq 2$$

where $C:=M^2/2-tr(Q)$.

Let $\lambda>0$ be any fixed real number. By the integration by parts formula, we have

$$\begin{aligned}\mathbb{E}^\nu e^{\lambda\tau_K^{(1)}} &= 1+\int_0^{+\infty}\lambda e^{\lambda t}\mathbb{P}_\nu\left(\tau_K^{(1)}>t\right)dt\\ &\leq e^{2\lambda}+\sum_{n\geq 2}\lambda e^{\lambda(n+1)}\mathbb{P}_\nu\left(\tau_K^{(1)}>n\right)\end{aligned}$$

$$\leq \quad e^{2\lambda}\left(1+\lambda\int_H e^{\lambda_0\|x\|_2^2}\nu(dx)\sum_{n\geq 2}e^{-n(\lambda_0 C-\lambda)}\right).$$

Now, by the definition (6.30) of K, we can choose M such that $\lambda_0 C - \lambda \geq 1$. Then, taking the supremum over $\{\nu = \delta_x,\ x \in K\}$, we get

$$\sup_{x\in K}\mathbb{E}^x e^{\lambda\tau_K^{(1)}} \leq e^{2\lambda}\left(1+\lambda e^{\lambda_0 M^2}\sum_{n\geq 2}e^{-n(\lambda_0 C-\lambda)}\right) < \infty$$

for $\forall x \in K$, $\|x\|_2 \leq \|\nabla x\|_2/\pi \leq M$. So the bound (6.28) holds true. We obtain (6.29) in the same way: since $\tau_K \leq \tau_K^{(1)}$, we have

$$\begin{aligned}
\sup_{\nu\in\mathcal{M}_{\lambda_0,L}}\mathbb{E}^\nu e^{\lambda\tau_K} &\leq \sup_{\nu\in\mathcal{M}_{\lambda_0,L}}\mathbb{E}^\nu e^{\lambda\tau_K^{(1)}}\\
&\leq e^{2\lambda}\left(1+\lambda L\sum_{n\geq 2}e^{-n(\lambda_0 C-\lambda)}\right)\\
&< \infty.
\end{aligned}$$

Proof of Theorem 6.1.1. At first the good uniform upper bound of large deviations, i.e., parts (i) and (iii) follows by Lemma 6.6.1 for its conditions (6.28) and (6.29) are verified above.

The uniform lower bound in part (ii) was established in Proposition 6.3.3.

The first claim in (6.6): "$J(\nu) < +\infty \implies \nu \ll \mu$" was proven in Lemma 6.3.1. We conclude the proof with the second claim in (6.6) that for $\nu \in M_1(H)$ with $J(\nu) < \infty$, $\nu(\|\nabla x\|_2^2) < \infty$. Indeed, denoting by $a \wedge b$ the minimum of two real numbers a and b, and for the function $V_n(x) := (\lambda_0\|\nabla x\|_2^2) \wedge n$ bounded and measurable on H, we have

$$\begin{aligned}
\nu(V_n) &\leq (\Lambda^0)^*(\nu) + \Lambda^0(V_n)\\
&\leq J(\nu) + \lambda_0 tr(Q)
\end{aligned}\tag{6.31}$$

where we have used the definitions (6.17), (6.18), the crucial estimate (6.26) and the fact that $(\Lambda^0)^* = J$. The conclusion follows by Fatou's lemma. □

6.7 Extension to some unbounded functionals

In this section we point out the fact that the estimate in Lemma 6.5.2 is sufficient to extend the LDP of Theorem 6.1.1, i.e. Corollary 6.1.2 for unbounded functionals and its consequences.

Proof of Corollary 6.1.2. For the function $f : H_0^1 \to \mathbb{B}$ measurable and bounded on balls, let us consider $f_n : H \to \mathbb{B}$ defined by

$$f_n(x) := \begin{cases} f(x), & \text{if } x \in H_0^1,\ \|\nabla x\|_2 \le n \\ 0 & \text{otherwise} \end{cases} \tag{6.32}$$

which is far from being continuous, but is measurable and bounded on H. Since $\nu \to \nu(f_n) = \int_{\mathbb{B}} z\nu(f_n \in dz)$ is continuous from $(M_1(H), \tau)$ to $\mathbb{B}$ by [12, Lemma 3.3.8], then $L_T(f_n)$ satisfies the LDP by Theorem 6.1.1 and the contraction principle.

Now by the approximation lemma in large deviations (see [12, p.37]), it remains to prove that for any $L > 0$

$$\lim_{n\to\infty} \sup_{\beta:J(\beta)\le L} \|\beta(f_n) - \beta(f)\|_{\mathbb{B}} = 0 \tag{6.33}$$

and for any $\delta > 0$,

$$\lim_{n\to\infty} \limsup_{T\to\infty} \frac{1}{T} \log \sup_{\nu\in\mathcal{M}_{\lambda_0,L}} \mathbb{P}_\nu \left(\|L_T(f - f_n)\|_{\mathbb{B}} > \delta\right) = -\infty. \tag{6.34}$$

Thanks to our condition (6.7) on f, we can construct a sequence $(\varepsilon(n))_n$ decreasing to 0 such that once $\|\nabla x\|_2 \ge n$,

$$\|f(x)\|_{\mathbb{B}} \le \varepsilon(n)\|\nabla x\|_2^2.$$

Denoting by 1_A the characteristic function of the set A, we have for any β satisfying $J(\beta) \le L$,

$$\begin{aligned} \|\beta(f_n) - \beta(f)\|_{\mathbb{B}} &= \|\beta\left(f 1_{\{\|\nabla X(s)\|_2 \ge n\}}\right)\|_{\mathbb{B}} \\ &\le \beta\left(\varepsilon(n)\|\nabla x\|_2^2 1_{\{\|\nabla X(s)\|_2\ge n\}}\right) \\ &\le \frac{\varepsilon(n)}{\lambda_0}\beta\left(\lambda_0\|\nabla x\|_2^2\right) \\ &\le \frac{\varepsilon(n)}{\lambda_0}\left(L + \lambda_0 tr(Q)\right) \end{aligned}$$

by using (6.31). Hence (6.33) follows.

Let us evaluate

$$\begin{aligned} \mathbb{P}_\nu\left(\|L_T(f - f_n)\|_{\mathbb{B}} > \delta\right) &= \mathbb{P}_\nu\left(\|\frac{1}{T}\int_0^T f(X_s) - f_n(X_s) ds\|_{\mathbb{B}} > \delta\right) \\ &\le \mathbb{P}_\nu\left(\frac{1}{T}\int_0^T \varepsilon(n)\|\nabla X(s)\|_2^2 1_{\{\|\nabla X(s)\|_2\ge n\}} ds > \delta\right) \\ &\le \mathbb{P}_\nu\left(\int_0^T \lambda_0\|\nabla X(s)\|_2^2 1_{\{\|\nabla X(s)\|_2\ge n\}} ds > \frac{\lambda_0 T\delta}{\varepsilon(n)}\right) \end{aligned}$$

$$\leq \exp\left(-\frac{\lambda_0 T\delta}{\varepsilon(n)}\right)\mathbb{E}^{\nu}\exp\left(\lambda_0\int_0^T \|\nabla X(s)\|_2^2 ds\right)$$

so that (6.34) is consequence of (6.27). □

Remerciements :

L'auteur remercie vivement Liming Wu pour les nombreuses discussions enrichissantes et l'attention qu'il a portée à ce travail.

Bibliography

[1] R. A. Adams, *Sobolev Spaces*, Academic Press New-York (1975).

[2] J.M Burgers, *The Nonlinear Diffusion Equation. Asymptotic Solutions and Statistical Problems*, Reidel (1970).

[3] C. Cardon-Weber, Large Deviations for a Burgers type SPDE, *Stoch. Proc. Appl.* **84**, *53-70 (1999).*

[4] G. Da Prato, A. Debussche, Differentiability of the transition semigroup of the Stochastic Burgers equation, and application to the corresponding Hamilton Jacobi equation, *Atti Accad. Naz Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei Mat. Appl.* **9**, *267-277 (1998).*

[5] G. Da Prato, A. Debussche, Dynamic Programming for the Stochastic Burgers equation, *Annali di Matematica pura ed applicata (IV), Vol. CLXXVIILI, 143-174 (2000).*

[6] G. Da Prato, D. Gatarek, Stochastic Burgers equation with correlated noise, *Stochastics and Stochastic Reports* **52**, *29-41 (1995).*

[7] G. Da Prato, A. Debussche and R. Temam, Stochastic Burgers equation, *Nonlinear Analysis* **1**, *389-402 (1994).*

[8] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge University Press (1992).

[9] G. Da Prato and J. Zabczyk, *Ergodicity for Infinite Dimensional Systems*, Cambridge University Press (1996).

[10] A. Dembo and O. Zeitouni, *Large Deviations Techniques and applications, 2nd edition*, Springer-Verlag (1998).

[11] A. Dermoune, Around the Stochastic Burgers equation, *Stoch. Analysis and Applic.,* **15(3)**, *295-311 (1997).*

[12] J.D. Deuschel and D.W. Stroock, *Large Deviations*, Pure Appl. Math., Vol. 137, Academic Press, San Diego (1989).

[13] M.D Donsker and S.R.S Varadhan, Asymptotic evalutation of certain Markov process expectations for large time, I-IV, *Comm. Pur. Appli. Math,* **28**, *1-47 and 279-301 (1975);* **29**, *389-461 (1976);* **36**, *183-212 (1983).*

[14] Weinan E, K. Khanin, A. Mazel and Ya. Sinai, Invariant measures for Burgers equation with stochastic forcing, *Ann. of Math.,* **151**, *877-960 (2000).*

[15] F. Flandoli, B. Maslowski, Ergodicity of the 2D Navier-Stokes Equation Under Random Perturbation, *Communications in Mathematical Physics,* **171**, *119-141 (1995).*

[16] B. Goldys, B. Maslowski, Exponential ergodicity for stochastic Burgers and 2D Navier-Stokes equation, *J. Funct. Anal.,* **226**, *No.1, 230-255 (2005).*

[17] B. Maslowski, On probability distributions of solutions of semilinear stochastic evolution equations, *Stochastics. Stoc. Rep.* **45**, *17-44 (1993)*

[18] S.P Meyn and R.L. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag (1993).

[19] E. Pardoux, *Equations aux dérivées partielles stochastiques non linéaires monotones*, Ph.D. thesis. Université Paris XI (1975).

[20] B.L. Rozovski, *Stochastic Evolution Systems: Linear theory and application to non linear filtering*, Kluver Academic Pub. (1990).

[21] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion, third edition*, Springer-Verlag (1999).

[22] L. Wu, Uniformly integrable operators and large deviations for Markov processes, *J. Funct. Anal.* **172**, *301-376 (2000).*

[23] L. Wu, Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems, *Stoch. Proc. Appl.* **91**, *205-238 (2001).*

[24] L. Wu, Essential spectral radius for Markov semigroups (I): discrete time case, *Proba. Theo. Rel. Fields ,* **128**, *255-321 (2004).*

Part III

Annexe

Chapitre 7

Compléments sur les grandes déviations de processus de Markov

On présente quelques résultats connus de Grandes Déviations pour la mesure d'occupation d'un processus de Markov fortement Feller (FF) et topologiquement irréductible (TI) qui sont cités, appliqués dans les chapitres 3, 4, 6 ou qui interviennent dans la preuve du Théorème 3.2.5.

Afin de faciliter la lecture de ce document, on reprend les notations de la section 3.2.3 : le processus de Markov X_t à trajectoires continues dans un espace polonais E est dénoté par

$$\left(\Theta, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, (X_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E}\right).$$

Son semigroupe de transition P_t possède les propriétés (FF) et (TI) et l'unique mesure invariante est notée μ. On renvoie aux références pour les résultats sous des conditions plus générales.

Dans un premier temps, nous rappelons une borne inférieure générale de grandes déviations (3.54) prouvée par Wu [7], qui permet d'obtenir la borne inférieure uniforme (3.4) si la borne supérieure (3.5) est valide comme on l'a vu dans le chapitre précédent.

Nous rappelons ensuite la procédure classique pour établir une borne supérieure faible (Gärtner-Ellis, voir [2]). Nous expliquons enfin comment les propriétés (FF) et (TI) ainsi que les conditions de retour de type hyper-exponentielles (3.7) permettent d'obtenir la borne supérieure uniforme (3.4) pour la topologie τ (voir [8]).

7.1 Borne inférieure de grandes déviations

Pour des chaînes de Markov irréductibles, un résultat sur la borne inférieure a été établi par De Acosta [1], puis étendu au cas des temps continus par Jain [4].

Nous présentons ici un cas particulier d'un résultat que Wu a prouvé dans le cadre des processus essentiellement irréductibles [6, 7].

Dans sa forme originale, ce résultat fait intervenir la restriction de l'entropie de Donsker Varadhan à sa composante sur μ introduite dans [5] et définie par

$$H_\mu(Q) := \begin{cases} H(Q) & \text{si } Q_0 := Q(X_0 \in \cdot) << \mu \\ \infty & \text{sinon} \end{cases}$$

où H est définie par (3.9). Cependant, sous les conditions (FF) et (TI), il n'y a pas de problème de transition de phase et on a en fait $H = H_\mu$ (voir chapitre 6). On rappelle que la topologie limite projective sur $M_1(\Theta)$ est définie par $\tau_p = \sigma\left(M_1(\Theta), \bigcup_{t\geq 0} b\mathcal{F}_t\right)$ où $b\mathcal{F}_t$ est l'ensemble des fonctions bornées sur Θ et mesurables pour la filtration contenant l'information jusqu'au temps t. On a le

Théorème 7.1.1. *Pour tout ouvert O dans $(M_1(\Theta), \tau_p)$,*

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}_x(R_t \in O) \geq -\inf_O H_\mu, \quad \mu - pp\ x \in H \tag{7.1}$$

où μ est l'unique mesure invariante.

Démonstration. Considérons un ouvert G dans $(M_1(\Theta), \tau_p)$ et soit Q un élément de G tel que $H_\mu(Q) < \infty$, sinon (7.1) serait immédiate. Au vu de la définition, Q est donc une mesure stationnaire, et pour tout $t > 0$, la mesure $Q|_{\mathcal{F}_t}$ est absolument continue par rapport à $\mathbb{P}_\mu|_{\mathcal{F}_t}$. Ainsi on a en particulier $Q_0 << \mu$. Un système fondamental de voisinages de Q pour la topologie τ_p sur $M_1(\Theta)$ est de la forme $N(Q)$ suivante : on sait qu'il existe $N \in \mathbb{N}$, un d-uplet de fonctions mesurables pour $\mathcal{F}_N$ et bornées $F := (F_1, F_2, ..., F_d)$, et un réel ε strictement positif tels que

$$N(Q) := \left\{ Q' \in M_1(\Theta) \text{ tels que } \left| \int F dQ' - \int F dQ \right|_d < \varepsilon \right\} \subset G \tag{7.2}$$

où $|\cdot|_d$ est la norme euclidienne dans $\mathbb{R}^d$.

En fait, pour démontrer la borne inférieure (7.1), il suffit d'établir que pour toute mesure stationnaire Q vérifiant $H_\mu(Q) < \infty$ et tout τ_p-voisinage $N(Q)$ de la forme précédente, on ait (caractère local de la borne inférieure de grandes déviations)

$$\liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}_x\left(R_t \in N(Q)\right) \geq -H_\mu(Q) \quad \mu - pp\ x \in H \tag{7.3}$$

La preuve se fait en trois étapes. On travaille d'abord dans un espace discret en temps, avec une mesure Q ergodique, on étend ensuite au cas où la mesure Q est quelconque, et enfin on passe en temps continu.

7.1.1 Temps discret, le cas Q ergodique et $\mu << Q_0$.

Nous supposons donc que Q est ergodique, et que l'on a $\mu \sim Q_0$. Dans cette section, nous considérons une chaîne de Markov indexée en temps discret (au lieu de notre processus initial indexé en temps continu) mais qui vit dans un espace plus grand.

Pour $s > 0$, soit $Y_n^{(s)} := X_{[ns,(n+1)s)}$, la chaîne $(Y_n^{(s)})_{n\in\mathbb{N}}$, qui sera aussi notée $(\omega_n)_{n\in\mathbb{N}}$ par abus, prend donc ses valeurs dans $F := C\left([0,s), E\right)$ l'ensemble des fonctions continues de $[0,s)$ dans E, et la loi de la chaine partant de x notée encore $\mathbb{P}_x$ est regardée comme une mesure sur disons $\Omega' := F^{\mathbb{N}}$. L'entropie de niveau 3 associée au processus $(Y_n)_n$ est notée $H^{(s)}(Q)$, et on a :

$$H^{(s)}(Q) = \mathbb{E}^{\bar{Q}} h_{\mathcal{F}_s^0}\left(\bar{Q}_{X(-\infty,0)}; \mathbb{P}_x\right)$$

Les objets (mesures, tribus...) définis précédemment s'adaptent naturellement à ce nouveau cadre. Par exemple, la tribu contenant l'information entre les temps m et n est $\tilde{\mathcal{F}}_n^m := \sigma\left(Y_k, m \leq k \leq n\right) = \mathcal{F}_{(n+1)s}^{ms}$. De plus, Q est une mesure ergodique sur Ω' et son unique extension stationnaire à l'espace $\bar{\Omega}' := F^{\mathbb{Z}}$, notée $\bar{Q}$, est donc aussi ergodique. Enfin on note ici $\bar{Q}_{\omega-}$ une version régulière de la distribution conditionnelle de $\bar{Q}$ sachant $\tilde{\mathcal{F}}_0^{-\infty}$.

On rappelle que dans le cas discret en temps, la mesure empirique du processus prend la forme $R_n^{(s)}(\omega) := \frac{1}{n}\sum_{k=0}^n \delta_{\theta_k\omega}$ où θ_k est un opérateur de décalage (*shift*) de k pas, i.e $\theta_k\omega_l := \omega_{k+l}$. La base de voisinage conserve la forme donnée par le côté gauche de (7.2), mais les fonctions F_i intervenant dans la définition de $N(Q)$ sont mesurables pour $\tilde{\mathcal{F}}_N$ (au lieu de $\mathcal{F}_N$).

Soit $\Psi(\omega)$ la densité de $\bar{Q}_{\omega-}$ par rapport à $\mathbb{P}_x$ sur $\tilde{\mathcal{F}}_1^1 = \sigma(\omega_1)$, dont l'existence est assurée par le fait que $H(Q) < \infty$. En itérant la décomposition

$$\begin{aligned}
\frac{d\bar{Q}_{\omega-}}{d\mathbb{P}_x}(\omega_1, ..., \omega_n) &= \frac{d\bar{Q}_{\omega-}}{d\mathbb{P}_x}(\omega_1, ..., \omega_{n-1})\ \frac{d\bar{Q}_{\omega-}}{d\mathbb{P}_x}(\omega_n|\omega_1, ..., \omega_{n-1}) \\
&= \frac{d\bar{Q}_{\omega-}}{d\mathbb{P}_x}(\omega_1, ..., \omega_{n-1})\ \Psi(\theta_{n-1}\omega)
\end{aligned}$$

on obtient, comme $\bar{Q}$ est stationnaire,

$$\frac{d\bar{Q}_{\omega-}}{d\mathbb{P}_x}|_{\tilde{\mathcal{F}}_n^0} = \exp\left(\sum_{k=0}^{n-1} \log\Psi(\theta_k\omega)\right).$$

Pour $\delta > 0$, on pose

$$W_n := \{\omega\ : R_n^{(s)}(\omega) \in N(Q)\} \quad \text{et} \quad D_{n,\delta} := \left\{\omega\ : \frac{1}{n}\sum_{k=0}^{n-1} \log\Psi(\theta_k\omega) \leq H^{(s)}(Q) + \delta\right\}.$$

Vu la forme de $N(Q)$, on sait que $W_n \in \tilde{\mathcal{F}}^0_{n+N}$. On voit de plus que $D_{n,\delta} \in \tilde{\mathcal{F}}^0_n$. On a alors $\bar{Q}$-p.s

$$\begin{aligned}\mathbb{P}_{w(0)}(W_n) &\geq \int_{W_n} \exp\left(-\sum_{k=0}^{n+N-1} \log \Psi(\theta_k \omega)\right) d\bar{Q}_{\omega-} \\ &\geq \exp\left[-(n+N)\left(H(Q)+\delta\right)\right] \bar{Q}_{\omega-}\left(W_n \bigcap D_{n+N,\delta}\right).\end{aligned}$$

Or, d'après le théorème de Fubini, si $\bar{Q}(A) = 1$, comme $\bar{Q}(A) = \int \bar{Q}_{\omega-}(A) d\bar{Q}(\omega)$, on a donc $\bar{Q}$-presque pour tout $\omega \in \Omega'$, $\bar{Q}_{\omega-}(A) = 1$. De plus d'après le théorème ergodique (notre hypothèse sur Q), on sait que $\bar{Q}$-presque pour tout $\omega \in \Omega'$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \log \Psi(\theta_k \omega) = \mathbb{E}^{\bar{Q}} \log \Psi = H^{(s)}(Q).$$

Ainsi on obtient $\bar{Q}$-presque pour tout $\omega \in \Omega'$,

$$\lim_{n\to\infty} \bar{Q}_{\omega-}\left(W_n \bigcap D_{n+N,\delta}\right) = 1$$

et donc comme on a supposé que $\mu << Q_0$, on a pour tout $\delta > 0$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x\left(R_n^{(s)} \in N(Q)\right) \geq -H^{(s)}(Q) - \delta \quad \mu - pp\ x \in H. \tag{7.4}$$

Cela suffit pour établir la borne locale (7.3), puisque δ est arbitraire.

7.1.2 Temps discret, le cas général

Pour étendre à une mesure Q quelconque, on utilise le lemme de construction suivant, établi dans le livre de Deuschel et Stroock [3] (lemme 4.4.9 p 165).

Lemme 7.1.2. *Soit Q telle que $H^{(s)}(Q) < \infty$ (et donc Q est une mesure stationnaire sur Ω'). Pour $d \geq 1$, soit*

$$Y_k^{(d)} := (Y_k, ..., Y_{k+d}) \in F^{d+1}, \quad k \geq 0 \quad et \quad \nu := Q\left(Y_k^{(d)} \in \cdot\right) \in M_1(F^{d+1})$$

alors

$$\inf_{Q' \in M_1^s(\Omega') \quad et \quad Q'(Y_0^{(d)} \in \cdot) = \nu} H^{(s)}(Q')$$

est atteint en une unique mesure $Q^{(d)} \in M_1^s(\Omega')$ caractérisée par

$$\begin{cases} Q^{(d)}(Y_k^{(d)} \in \cdot) = \nu, \quad \forall k \in \mathbb{N} \\ Q^{(d)}(Y_{n+1} \in \cdot | \mathcal{F}_n) = Q(Y_{n+1} \in \cdot | Y_{n-d+1}, ..., Y_n), \quad \forall n \geq d \end{cases}.$$

Le processus $(Y_k^{(d)})_{k\geq 0}$ est ainsi une chaîne de Markov sous $Q^{(d)}$: on dit que $Q^{(d)}$ une d-Markov modification de Q.

Pour prouver la borne inférieure (7.3) pour Q non ergodique, il suffit de construire une suite $(Q^m)_{m\geq 1}$ telle que

(a) Q^m est ergodique et $Q^m \sim \mathbb{P}_\mu$ sur $\tilde{\mathcal{F}}_n^0,\ \forall n \in \mathbb{N}$

(b) $Q^m \to Q$ en topologie τ_p et $\lim_{m\to\infty} H^{(s)}(Q^m) \leq H^{(s)}(Q)$.

Voici une construction rendue possible par le lemme ci-dessus : pour tout $m \geq 1$, on définit

$$Q^m := \left((1-m^{-1})Q + m^{-1}\mathbb{P}_\mu\right)^{(m)}$$

de sorte que l'on ait $Q^m \to Q$ en topologie τ_p et

$$\begin{array}{rcl} H^{(s)}(Q^m) & \leq & H^{(s)}\left((1-m^{-1})Q + m^{-1}\mathbb{P}_\mu\right) \\ & = & (1-m^{-1})H^{(s)}(Q) + m^{-1}H^{(s)}(\mathbb{P}_\mu). \end{array}$$

Comme le terme de droite admet $H^{(s)}(Q)$ pour limite, on a obtenu le (b).

Il ne reste plus qu'à vérifier le (a). Montrer l'ergodicité de Q^m revient à montrer l'ergodicité sous Q^m du processus de Markov $(Y_k^{(m)})_{m\geq 0}$ défini dans le lemme ci dessus. Le résultat suivant établi par Wu [7] (lemme 3.5 p314) peut être utilisé.

Lemme 7.1.3. *Soit P un noyau de transition admettant une mesure invariante μ. Si P est μ-essentiellement irréductible, alors μ est ergodique.*

Il s'agit alors de remarquer que le processus de Markov $(Y_k^{(m)})_{k\geq 0}$ est Q^m-essentiellement irréductible. On sait que sous $\mathbb{P}_\mu$, le processus X_t est μ-essentiellement irréductible, ce qui implique que le processus $(Y_k^{(s)})_{k\geq 0}$ est irréductible pour la mesure $\mathbb{P}_\mu$ restreinte aux informations de $\mathcal{F}_s^0$. Par suite, le processus $(Y_k^{(m)})_{k\geq 0}$ est irréductible pour la mesure $\mathbb{P}_\mu$ restreinte aux informations de $\tilde{\mathcal{F}}_m^0$.

Comme pour tout $n \in \mathbb{N}$, on a $Q << \mathbb{P}_\mu$ sur $\tilde{F}_n^0$, on sait que pour tout $m > 0$, $(1-m^{-1})Q + m^{-1}\mathbb{P}_\mu \ \sim\ \mathbb{P}_\mu$ sur $\tilde{F}_n^0$, pour tout $n \in \mathbb{N}$. Il en résulte que $(Y_k^{(m)})_{k\geq 0}$ est Q^m-essentiellement irréductible et que $Q^m \sim \mathbb{P}_\mu$ sur $\tilde{\mathcal{F}}_n^0,\ \forall n \in \mathbb{N}$.

Ceci achève la démonstration de la borne (7.1) pour le processus en temps discret $(Y_n^{(s)})_{n\geq 0}$.

7.1.3 Passage en temps continu

Dans ce paragraphe, on note encore $H^{(s)}(Q)$ l'entropie de niveau 3 de Donsker Varadhan et $R_n^{(s)}$ le processus empirique de niveau 3 associés à la chaîne de Markov $Y_n^{(s)}$ définie ci-dessus. On rappelle que si $H(Q) < \infty$, on a

$$H(Q) = \frac{1}{s}\mathbb{E}^{\bar{Q}} h_{\mathcal{F}_s^0}\left(\bar{Q}_{\omega(-\infty,0)}; \mathbb{P}_{w(0)}\right) = \frac{H^{(s)}(Q)}{s}$$

On *approche* le voisinage $N(Q)$ considéré dans (7.3), avec un ouvert $\tilde{N}(Q)$ concernant la chaîne $Y_n^{(s)}$ défini par

$$\tilde{N}(Q) := \left\{ Q' \text{ tels que } \left| \int \tilde{F} dQ' - \int \tilde{F} dQ \right|_d < \frac{\varepsilon}{2} \right\}$$

où $\tilde{F}(\omega) = \frac{1}{s}\int_0^s F(\theta_u\omega)du$. Ainsi, pour tout $t \geq s$, s'écrivant donc $t = ns + r$ avec $0 \leq r < s$ et $n \in \mathbb{N}$, on a

$$\begin{aligned} \left| \int F dR_t - \int \tilde{F} dR_n^{(s)} \right|_d &= \left| \frac{1}{t}\int_0^t F(\theta_u\omega)du - \frac{1}{ns}\int_0^{ns} F(\theta_u\omega)du \right|_d \\ &\leq \frac{2s}{t}||F||_\infty. \end{aligned}$$

Pour t suffisamment grand, le membre de droite est majoré par $\frac{\varepsilon}{2}$ ce qui implique l'inclusion

$$\left[R_n^{(s)} \in \tilde{N}(Q) \right] \subset [R_t \in N(Q)] .$$

Grâce au cas discret, on a alors μ-presque pour tout x dans E,

$$\begin{aligned} \liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}_x \left(R_t \in N(Q) \right) &\geq \liminf_{n\to\infty} \frac{1}{ns} \log \mathbb{P}_x \left(R_n^{(s)} \in \tilde{N}(Q) \right) \\ &\geq -\frac{1}{s} H^{(s)}(Q) \\ &\geq -H(Q) \end{aligned}$$

ce qui achève la preuve.

□

7.2 Borne supérieure de grandes déviations

Dans cette section nous rappelons d'abord la méthode de Cramer-Gärtner-Ellis (voir [2, 3]) qui permet d'obtenir une borne supérieure faible. Puis nous donnons les conditions permettant d'obtenir la borne supérieure uniforme (3.5) pour la topologie τ, et nous montrons que ces conditions sont impliquées par les hypothèses (3.7) du critère.

7.2.1 Fonctionnelles de Cramer et borne supérieure faible

La fonctionnelle de Cramer uniforme sur un ensemble de mesures initiales $\mathcal{M}$ de $M_1(E)$ est définie pour $V \in b\mathcal{B}(E)$ par

$$\Lambda(V|\mathcal{M}) := \limsup_{t\to\infty} \frac{1}{t} \log \sup_{\nu\in\mathcal{M}} \mathbb{E}^\nu \exp\left(tL_t(V) \right) .$$

Il s'agit de l'une des versions introduites et étudiées dans [3]. En effet, on peut aussi considérer

$$\begin{aligned}
\Lambda(V|x) &:= \Lambda(V|\{\delta_x\}) = \limsup_{t\to\infty} \frac{1}{t} \log \mathbb{E}^x \exp\left(tL_t(V)\right) \\
\Lambda^\infty(V) &:= \Lambda(V|\{\delta_x \ ; \ x \in E\}) = \limsup_{t\to\infty} \frac{1}{t} \log \sup_{x\in E} \mathbb{E}^x \exp\left(tL_t(V)\right) \\
\Lambda^0(V) &:= \sup_{x\in E} \Lambda(V|x) = \sup_{x\in E} \limsup_{t\to\infty} \frac{1}{t} \log \mathbb{E}^x \exp\left(tL_t(V)\right)
\end{aligned}$$

où $V \in b\mathcal{B}(E)$. Les plus connues, $\Lambda^0(V)$ et $\Lambda^\infty(V)$ sont respectivement les fonctionnelles ponctuelles et uniformes de Cramer. Pour $\Lambda : b\mathcal{B}(E) \to \mathbb{R}$ l'une de ces fonctionnelles, sa transformée de Legendre est donnée par

$$\begin{aligned}
\Lambda_w^*(\nu) &= \sup_{V\in C_b(E)} \left(\int_E V d\nu - \Lambda(V) \right), \quad \forall \nu \in M_1(E) \\
\text{ou} \quad \Lambda^*(\nu) &= \sup_{V\in b\mathcal{B}(E)} \left(\int_E V d\nu - \Lambda(V) \right), \quad \forall \nu \in M_1(E)
\end{aligned}$$

selon la topologie considérée sur $M_1(E)$. Dans cette section, $(M_1(E), w)$ est équipé de la topologie faible de la convergence contre les fonctions continues et bornées.

La fonctionnelle de Cramer uniforme sur l'ensemble $\mathcal{M}$ donnée par

$$\Lambda(V|\mathcal{M}) = \limsup_{T\to\infty} \frac{1}{T} \log \sup_{\nu\in\mathcal{M}} \mathbb{E}^\nu \left(e^{\int_0^T V(X_s)ds} \right), \qquad \forall V \in C_b(E)$$

est une fonction convexe, comme limite supérieure de fonctions convexes. La transformée de Legendre associée est

$$\Lambda_{\mathcal{M}}^*(\mu) = \sup_{V\in C_b(E)} \mu(V) - \Lambda(V|\mathcal{M}), \qquad \forall \mu \in M_1(E). \tag{7.5}$$

Le résultat suivant est un cas particulier du théorème de Gärtner-Ellis, très largement connu dans la théorie des grandes déviations.

Théorème 7.2.1. *Pour tout ensemble compact K dans $(M_1(E), w)$ muni de la topologie faible, et pour tout réel strictement positif ε petit, il existe un voisinage mesurable $N(K, \varepsilon)$ de K dans $(M_1(E), w)$ tel que*

$$\begin{aligned}
&\limsup_{T\to\infty} \frac{1}{T} \log \sup_{\nu\in\mathcal{M}} \mathbb{P}^\nu \left(L_T \in N(K,\varepsilon)\right) \\
\le \ &\begin{cases} -\inf_{\nu\in K} \Lambda_{\mathcal{M}}^*(\nu) + \varepsilon & \textit{si} \quad \inf_{\nu\in K} \Lambda_{\mathcal{M}}^*(\nu) < \infty \\ -\frac{1}{\varepsilon} & \textit{sinon.} \end{cases}
\end{aligned}$$

Démonstration. Soit un compact K de $(M_1(E), w)$ et $\varepsilon > 0$ un réel. Pour $\mu \in K$, vu l'expression (7.5) donnant la transformée de Legendre, on peut trouver $V_\mu \in C_b(E)$ telle que

$$\mu(V_\mu) - \Lambda(V_\mu|\mathcal{M}) \geq \min\{\Lambda^*_{\mathcal{M}}(\mu) - \varepsilon \ , \ \frac{1}{\varepsilon}\}.$$

Comme la fonction de $M_1(E)$ dans $\mathbb{R}$ qui à ν fait correspondre $\nu(V_\mu)$ est continue, il existe un voisinage A_μ de μ dans $(M_1(E), w)$ tel que

$$\inf_{\lambda \in A_\mu} \{\lambda(V_\mu) - \mu(V_\mu)\} \geq -\varepsilon.$$

Maintenant, pour toute mesure initiale ν dans $\mathcal{M}$, et pour toute fonction $V \in C_b(E)$, on a d'après l'inégalité de Chebycheff

$$\begin{aligned}
\mathbb{P}^\nu(L_T \in A_\mu) &= \int_{A_\mu} \mathbb{P}^\nu(L_T \in dx) \\
&\leq \int_{A_\mu} \frac{e^{x(V)}}{e^{\inf_{\lambda \in A_\mu} \lambda(V)}} \mathbb{P}^\nu(L_T \in dx) \\
&\leq \int_{M_1(E)} \frac{e^{x(V)}}{e^{\inf_{\lambda \in A_\mu} \lambda(V)}} \mathbb{P}^\nu(L_T \in dx) \\
&= \mathbb{E}^\nu \left(e^{L_T(V) - \inf_{\lambda \in A_\mu} \lambda(V)} \right) \\
&= \mathbb{E}^\nu \left(e^{L_T(V) - \mu(V)} \right) e^{-\inf_{\lambda \in A_\mu} \{\lambda(V) - \mu(V)\}}.
\end{aligned}$$

Si on applique l'estimation obtenue à $V = TV_\mu$, on obtient

$$\mathbb{P}^\nu(L_T \in A_\mu) \leq \mathbb{E}^\nu \left(e^{\int_0^T V_\mu(X_s) ds - \mu(TV_\mu)} \right) e^{-\inf_{\lambda \in A_\mu} \{\lambda(TV_\mu) - \mu(TV_\mu)\}}$$

puis, en prenant d'abord le supremum sur les mesures initiales de $\mathcal{M}$, et ensuite le logarithme, on trouve

$$\begin{aligned}
& \frac{1}{T} \log \sup_{\nu \in \mathcal{M}} \mathbb{P}^\nu(L_T \in A_\mu) \\
\leq \ & \frac{1}{T} \log \sup_{\nu \in \mathcal{M}} \mathbb{E}^\nu \left(e^{\int_0^T V_\mu(X_s) ds} \right) - \mu(V_\mu) - \inf_{\lambda \in A_\mu} \{\lambda(V) - \mu(V_\mu)\} \\
\leq \ & \frac{1}{T} \log \sup_{\nu \in \mathcal{M}} \mathbb{E}^\nu \left(e^{\int_0^T V_\mu(X_s) ds} \right) - \mu(V_\mu) + \varepsilon.
\end{aligned}$$

La compacité de K nous permet alors d'extraire un sous recouvrement fini :

$$N(K, \varepsilon) := \bigcup_{i=1}^{i=N} A_{\mu_i}.$$

On a donc pour ce voisinage mesurable de K,

$$\begin{aligned}
\frac{1}{T}\log\sup_{\nu\in\mathcal{M}}\mathbb{P}^\nu(L_T\in K) &\le \frac{1}{T}\log\sup_{\nu\in\mathcal{M}}\mathbb{P}^\nu\left(L_T\in N(K,\varepsilon)\right)\\
&\le \frac{1}{T}\log\sum_{i=1}^{i=N}\sup_{\nu\in\mathcal{M}}\mathbb{P}^\nu\left(L_T\in A_{\mu_i}\right)\\
&\le \frac{1}{T}\left[\log N+\max_{i=1}^{i=N}\log\sup_{\nu\in\mathcal{M}}\mathbb{P}^\nu\left(L_T\in A_{\mu_i}\right)\right]\\
&\le \frac{\log(N)}{T}+\varepsilon-\min_{i=1}^{i=N}\left[\mu_i(V_{\mu_i})-\frac{1}{T}\log\sup_{\nu\in\mathcal{M}}\mathbb{E}^\nu\left(e^{\int_0^T V_{\mu_i}(X_s)ds}\right)\right].
\end{aligned}$$

En prenant la limite supérieure lorsque T tend vers l'infini, on a

$$\begin{aligned}
\limsup_{T\to\infty}\frac{1}{T}\log\sup_{\nu\in\mathcal{M}}\mathbb{P}^\nu\left(L_T\in N(K,\varepsilon)\right) &\le -\min_{i=1}^{i=N}\left[\mu_i(V_{\mu_i})-\Lambda(V_{\mu_i}|\mathcal{M})\right]+\varepsilon\\
&\le -\min\left\{\min_{i=1}^{i=N}[\Lambda^*_{\mathcal{M}}(\mu_i)-\varepsilon]\ \ ;\ \ \frac{1}{\varepsilon}\right\}+\varepsilon\\
&\le -\min\left\{[\inf_{\nu\in K}\Lambda^*_{\mathcal{M}}(\nu)-2\varepsilon]\ \ ;\ \ \frac{1}{\varepsilon}-\varepsilon\right\}
\end{aligned}$$

en faisant l'infimum pour μ_i dans le compact K à gauche. Ce qui achève la preuve, quitte à changer en 2ε en ε en remarquant que $1/\varepsilon-\varepsilon\ge 1/2\varepsilon$ dès que $\varepsilon\le 1/4$. $\square$

La situation est encore plus intéressante lorsque $\mathcal{M}=\mathcal{M}_{\lambda_0,L}$ avec $\{\delta_x\}_{x\in E}\subset\cup_{L>0}\mathcal{M}_{\lambda_0,L}$, comme c'est le cas pour les modèles étudiés dans cette thèse. En effet, on a alors

$$\Lambda^0(V)\le\sup_{L>0}\Lambda(V|\mathcal{M}_{\lambda_0,L})\le\Lambda^\infty(V),\qquad\forall V\in b\mathcal{B}(E).$$

De plus, comme (P_t) possède la propriété de Feller, on a d'après [7, Proposition B.13],

$$\left(\Lambda^0\right)^*(\nu)=\left(\Lambda^0\right)^*_w(\nu)=\left(\Lambda^\infty\right)^*(\nu)=\left(\Lambda^\infty\right)^*_w(\nu)=J(\nu),\ \ \forall\nu\in M_1(E).$$

On en déduit que J est s.c.i, que

$$\begin{aligned}
&\sup_{V\in b\mathcal{B}(E)}\left(\int_E Vd\nu-\sup_{L>0}\Lambda(V|\mathcal{M}_{\lambda_0,L})\right)\\
&=\sup_{V\in C_b(E)}\left(\int_E Vd\nu-\sup_{L>0}\Lambda(V|\mathcal{M}_{\lambda_0,L})\right)=J(\nu),\ \forall\nu\in M_1(E).
\end{aligned}$$

et donc que $J\le\Lambda^*_{\mathcal{M}_{\lambda_0,L}}$ pour tout J et tout λ_0. La borne supérieure faible du théorème précédent est alors valide pour la fonction de taux J.

En fait les conditions (3.7) impliquent aisément la propriété de tension exponentielle qui permet d'obtenir le résultat pour tous les ensembles de $(M_1(H), w)$ fermés pour la topologie faible.

Le point restant est donc l'obtention de la borne supérieure pour la topologie τ.

7.2.2 De la topologie faible à la topologie τ

Des conditions suffisantes pour ce passage ont été étudiées entre autres par de Acosta, Léonard et Wu (voir [1, 5, 7]). D'après le théorème B.5 de [7], la bonne borne supérieure (pour la topologie τ) relativement à est $\Lambda^*_{\mathcal{M}}$ est équivalente à la décroissance vers 0 de la fonctionnelle de Cramer associée $\Lambda(f_n \mid \mathcal{M})$ pour toute suite de fonctions mesurables et bornées $(f_n)_n \subset b\mathcal{B}(E)$ qui décroît vers 0 en tout point $x \in E$. Pour compléter la preuve de la borne supérieure (3.5), et donc de notre Théorème 3.2.5, comme $J \le \Lambda^*_{\mathcal{M}_{\lambda_0,L}}$, il reste donc à établir le

Lemme 7.2.2. *Soit X_t un processus de Markov fortement Feller et topologiquement irréductible sur E qui vérifie les conditions de récurrence hyper-exponentielle (3.7).*

Alors, pour toute suite de fonctions $(f_n)_n \subset b\mathcal{B}(E)$ qui décroît ponctuellement vers 0*, on a*

$$\limsup_{n\to\infty} \Lambda(f_n \mid \mathcal{M}) := \limsup_{n\to\infty} \limsup_{T\to\infty} \frac{1}{T} \log \sup_{\nu\in\mathcal{M}} \mathbb{E}^\nu e^{\int_0^T f_n(X_s)ds} \;=\; 0.$$

Démonstration. On reprend ici la preuve du théorème 2.1 dans [8]. Soit $(f_n)_n$ une suite de fonctions mesurables bornées sur E telle que $f_n(x)$ décroît vers 0 pour tout x de E.

Soit $M_0 := 2\sup_E |f_0(z)|$, $M \ge M_0 + 1$ et K un compact tel que les conditions (3.7) soient vérifiées pour $\lambda = M$. On définit les temps de retour successifs en K par

$$\begin{aligned} \tau_K^0 &:= \tau_K \\ \tau_K^{m+1} &:= \inf\{t \ge \tau_k^m + 1 \text{ t.q. } X_t \in K\} \;\; ; \;\; \text{pour } m \ge 0. \end{aligned}$$

Soit $[T]$ la partie entière de T, et $\xi_k(n)$ la variable aléatoire réelle définie par

$$\xi_k(n) := \int_{\tau_K^k}^{\tau_K^{k+1}} f_n(X_s)ds.$$

Comme $\tau_K^{[T]+1} \ge [T] + 1 \ge T$, on a

$$\int_0^T f_n(X_s)ds \le \int_0^{\tau_K^1} f_n(X_s)ds + \sum_{k=1}^{[T]} \xi_k(n)$$

Par définition de M, et en utilisant la propriété de Markov on obtient

$$\begin{aligned}
\sup_{\nu\in\mathcal{M}} \mathbb{E}^\nu e^{\int_0^{\tau_K^1} f_n(X_s)ds} &\leq \sup_{\nu\in\mathcal{M}} \mathbb{E}^\nu e^{M\tau_K^1} \\
&\leq \left(\sup_{\nu\in\mathcal{M}} \mathbb{E}^\nu e^{M\tau_K}\right)\left(\sup_{z\in K} \mathbb{E}^z e^{M\tau_K(1)}\right) \\
&< \infty
\end{aligned}$$

puisque K satisfait à (3.7) pour $\lambda = M$. On a donc, en notant $m := [T]$,

$$\Lambda(f_n \mid \mathcal{M}) := \limsup_{T\to\infty} \frac{1}{T} \log \sup_{\nu\in\mathcal{M}} \mathbb{E}^\nu e^{\int_0^T f_n(X_s)} \leq \limsup_{m\to\infty} \frac{1}{m} \log \sup_{z\in K} \mathbb{E}^z e^{\sum_{k=1}^m \xi_k(n)}.$$

Pour estimer cette expression, on remarque que la propriété de Markov, et l'inégalité élémentaire $e^{a+b} \leq \frac{1}{2}(e^{2a} + e^{2b})$ impliquent que

$$\begin{aligned}
\mathbb{E}^z e^{\sum_{k=1}^{2m} \xi_k(n)} &\leq \frac{1}{2}\left(\mathbb{E}^z e^{\sum_{j=1}^m 2\xi_{2j-1}(n)} + \mathbb{E}^z e^{\sum_{j=1}^m 2\xi_{2j}(n)}\right) \\
&\leq \frac{1}{2}\left[\left(\sup_{z\in K} \mathbb{E}^z e^{2\xi_1(n)}\right)^m + \left(\sup_{z\in K} \mathbb{E}^z e^{2\xi_1(n)}\right)^m\right] \\
&\leq \left(\sup_{z\in K} \mathbb{E}^z e^{2\xi_1(n)}\right)^m.
\end{aligned}$$

En utilisant cette quantité dans l'inégalité précédente, on obtient maintenant pour $N > 1$ fixé,

$$\begin{aligned}
\Lambda(f_n|\mathcal{M}) &\leq \frac{1}{2} \log \left(\sup_{z\in K} \mathbb{E}^z e^{\int_{\tau_K^1}^{\tau_K^2} 2f_n(X_s)ds}\right) \\
&\leq \frac{1}{2} \log \left(\sup_{z\in K} \mathbb{E}^z e^{\int_{\tau_K^1}^{\tau_K^1+N} 2f_n(X_s)ds} + \sup_{z\in K} \mathbb{E}^z \mathbb{1}_{[\tau_K^1+N\leq\tau_K^2]} e^{\int_{\tau_K^1+N}^{\tau_K^2} 2f_n(X_s)ds}\right) \\
&\leq \frac{1}{2} \log \left(\sup_{z\in K} \mathbb{E}^z e^{\int_{\tau_K^1}^{\tau_K^1+N} 2f_n(X_s)ds} + \sup_{z\in K} \mathbb{E}^z \mathbb{1}_{[\tau_K^1+N\leq\tau_K^2]} e^{M_0(\tau_K^2-\tau_K^1)}\right).
\end{aligned}$$

En remarquant que le choix $\lambda = M$ supérieur à $M_0 + 1$ dans l'hypothèse (3.7) implique l'uniforme intégrabilité de la famille $\{e^{M_0\tau_K(1)}, \mathbb{P}_z\}_{z\in K}$, on a en particulier

$$\begin{aligned}
\varepsilon(N) &:= \sup_{z\in K} \mathbb{E}^z \left(\mathbb{1}_{[\tau_K^1+N\leq\tau_K^2]} e^{M_0(\tau_K^2-\tau_K^1)}\right) \\
&= \sup_{z\in K} \mathbb{E}^z \left(\mathbb{1}_{[\tau_K(1)>N]} e^{M_0\tau_K(1)}\right) \\
&\to 0 \text{ quand } N \to \infty.
\end{aligned}$$

De plus, pour $z \in K$, on a $\tau_K^1 = \tau_K^{(1)}$ et donc pour $N > 1$ fixé,

$$\mathbb{E}^z e^{\int_{\tau_K^1}^{\tau_K^1+N} 2f_n(X_s)ds} = P_1 F_n(z)$$

où $F_n(\cdot) := \mathbb{E}^{\cdot} e^{\int_{\tau_K}^{\tau_K+N} 2f_n(X_s)ds}$. Maintenant, comme $F_n(\cdot) \leq e^{MN}$, on obtient pour tout $z \in K$, la décroissance de $P_1F_n(z)$ vers 1 par convergence dominée. Comme P_1 possède la propriété forte de Feller, P_1F_n est donc une suite de fonctions continues qui décroît vers 1 sur le compact K. Le théorème de Dini nous donne donc la convergence uniforme

$$\lim_{n\to\infty} \sup_{z\in K} |P_1F_n(z) - 1| \to 0.$$

En rassemblant ces convergences on obtient pour tout $N > 1$,

$$\limsup_{n\to\infty} \Lambda(f_n|\mathcal{M}) \leq \frac{1}{2}\log(1+\varepsilon(N)).$$

On termine donc la démonstration, et la rédaction de ce document (ouf), en prenant la limite $N \to \infty$. □

Bibliographie

[1] A. de Acosta, Large Deviations for empirical measures of Markov Chains, *J. Theoret. probab* **3**, *395-431 (1990).*

[2] A. Dembo and O. Zeitouni, *Large Deviations Techniques and applications, 2nd edition*, Springer-Verlag (1998).

[3] J.D. Deuschel and D.W. Stroock, *Large Deviations*, Pure Appl. Math., Vol. 137, Academic Press, San Diego (1989).

[4] N.C Jain, Large Deviations for additive functionnals of Markov processes. *Ann. Proba.* **17(3)**, *1073-1098 (1990).*

[5] L. Wu, Grandes Déviations pour les processus de Markov essentiellement irréductibles. II temps continu. *CRAS I,* **314**, *941-946 (1992).*

[6] L. Wu, *Habilitation à diriger des recherches*, Laboratoire de Probabilités, Université Paris VI (1993).

[7] L. Wu, Uniformly integrable operators and large deviations for Markov processes, *J. Funct. Anal.* **172**, *301-376 (2000).*

[8] L. Wu, Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems, *Stoch. Proc. Appl.* **91**, *205-238 (2001).*

RESUME

On s'intéresse dans cette thèse au comportement ergodique de certains systèmes dynamiques.

Dans la première partie, on établit une inégalité de Sobolev logarithmique pour la loi d'un mouvement Brownien avec dérive X_t, et plus généralement de certaines diffusions elliptiques, sur l'espace des trajectoires muni d'une métrique L^2. Ce résultat implique des propriétés de concentration intéressantes pour le comportement en temps grands des fonctionnelles de la trajectoire du type $1/\sqrt{T}\int_0^T f(X_t)dt$.

Dans la seconde partie, on prouve un principe de grandes déviations pour la mesure empirique des équations de Burgers et de Navier-Stokes stochastiques. Ce principe décrit la convergence exponentielle vers la mesure d'équilibre du système, dont l'unicité est assurée par les conditions de non dégénérescence imposées sur la perturbation aléatoire.

ABSTRACT

This thesis is devoted to the study of the ergodic behavior of certain dynamical systems.

In the first part, we establish the logarithmic Sobolev inequality for the Brownian motion with drift X_t, and more generally for elliptic diffusions, on the path space equipped with the metric L^2. This inequality provides us interesting concentration properties for the large time behavior of additive functionals such as $1/\sqrt{T}\int_0^T f(X_t)dt$.

In the second part, one specifies the ergodic behavior of the stochastic Burgers and Navier-Stokes equations by giving a large deviation principle for empirical measure for large time. It describes the exponential convergence to the unique equilibrium state, when the random force is sufficiently non-degenerate.